THEORIES OF CHANGE IN REALITY

STRENGTHS, LIMITATIONS, AND FUTURE DIRECTIONS

Edited by
Andrew Koleros, Marie-Hélène Adrien,
and Tony Tyrrell

'This is a welcome contribution to inquire into the uses of theories of change (ToC), that are nowadays ubiquitous. The authors bravely criticize a widespread approach that speaks ToC but in fact reproduces old top-down, bureaucratic, 'colonial', project-bounded attributes. Considering a ToC approach as an inter-active process (dynamic, participatory, and versatile) as opposed to the framing of a product, and inquiring into the many theoretical and practical aspects of it, the book is a precious guide for the evaluation of policies and programmes that address the current challenges of ecological disasters, social inequalities, conflict and violence'.

Nicoletta Stame, *Sapienza University of Rome, Italy*

'In order to understand how policy interventions work and bring about change, theories of change have become the go-to frameworks for evaluators, decision makers and other stakeholders alike. This book is both timely and insightful in its comprehensive and critical treatment of the subject. It brings together a sizable, well-curated collection of short essays by a diverse group of experts. I highly recommend this book to anyone who seeks to better understand how to make sense of policy interventions in their contexts with limited data, resources, and time. Because let us not forget: not only in evaluation but in policy design and implementation more broadly, there is nothing as practical as a good theory'.

Jos Vaessen, *Evaluation Adviser, Independent Evaluation Group, World Bank*

Theories of Change in Reality

For over 50 years, evaluators have used theories of change to articulate the causal logic underpinning how an intervention is intended to bring about a desired change. From its origins in programme evaluation, the approach has been adopted more widely for purposes from program design to program management. As theories of change continue to be used for multiple purposes, it is an opportune moment for the evaluation community—where the approach originated—to provide their perspective on the strengths and limitations of the approach and its future directions. To provide these perspectives, we asked nearly 30 of the world's leading evaluators and programme theorists to provide a short essay on the past, present, and future of theories of change. This book presents their insights organized into five main themes: the use of theories of change in broader public policy contexts; using theories of change to establish causality; developing theories of change reflective of multiple stakeholder perspectives; using theories of change to understand wider societal change processes; and applying theories of change approaches for multiple purposes. By sharing these diverse perspectives, the book aims to provide both evaluators and emerging program theorists with critical perspectives to inform future practice.

Andrew Koleros is a Principal Researcher at Mathematica with 20 years of experience in designing and delivering mixed-methods evaluations and programme monitoring, evaluation and learning systems for both small-scale and large-scale social and economic development projects. He brings particular expertise in using theory-based approaches that integrate complexity concepts and systems thinking into programme and evaluation design processes to address complex societal problems from safe and affordable housing to youth employment to advancing health equity. He obtained his master's in public health from the Tulane School of Public Health and Tropical Medicine. He has published on his work in designing theory-based evaluation approaches in the *American Journal of Evaluation*, the *Canadian Journal of Programme Evaluation* and in multiple practitioner settings. He is a member of the American Evaluation Association, the European Evaluation Society, and the Canadian Evaluation Society.

Marie-Hélène Adrien is a senior evaluation practitioner and researcher with 35 years of experience in evaluating the performance of programmes and institutions financed by bilateral and multilateral agencies and philanthropies around the globe. Dr. Adrien is a former CEO of the consulting firm Universalia Management Group in Montreal and was the President of the International Development Evaluation Association (IDEAS), as well as a professor of practice at the International Development Studies Program of McGill University.

Tony Tyrrell is an independent consultant with more than 25 years' experience in evaluation and related fields. Tony's early experience in evaluation was with the European Social Fund Evaluation Unit where he produced influential evaluations on various subjects including early school leaving and local development. He later worked with a number of private consulting firms on strategic, policy, and programme evaluation and on performance management. More recently he spent six years with the Independent Evaluation Group at the World Bank Group (WBG). Tony now works as an independent consultant with various clients including the WBG, the Asian Development Bank, various governments, and private consultancies. He holds an MSc in Management (Organisation Behavior), and an MA in English Literature. Tony was co-editor (with Burt Perrin) of another book in this series, *Changing Bureaucracies* (2021).

Comparative Policy Evaluation
Edited by Ray C. Rist

The Comparative Policy Evaluation series is an interdisciplinary and internationally focused set of books that embodies within it a strong emphasis on comparative analyses of governance issues—drawing from all continents and many different nation states. The lens through which these policy initiatives are viewed and reviewed is that of evaluation. These evaluation assessments are done mainly from the perspectives of sociology, anthropology, economics, policy science, auditing, law, and human rights. The books also provide a strong longitudinal perspective on the evolution of the policy issues being analyzed.

Theories of Change in Reality

Strengths, Limitations, and Future Directions

**Edited by Andrew Koleros,
Marie-Hélène Adrien, and Tony Tyrrell**

Contents

Acknowledgements

List of contributors

Jens Andersson is a Swedish national with more than 20 years' experience leading strategy development, monitoring, and evaluation at the intersection of business, philanthropy and development cooperation. As an independent consultant, he has worked with organizations such as the World Bank, WTO, WHO, EU, DFID, Swedish Ministry for Foreign Affairs, Sida, and Norad. Since 2019, he has been responsible for monitoring, evaluation, and learning at IKEA Social Entrepreneurship, based in The Netherlands. Andersson's work also involves providing advice on social impact and supporting theory of change formulation for various sustainability initiatives within and outside IKEA. He has a PhD in Economic History from Lund University.

Kathy Chau Rohn is a postdoctoral research associate at the University of Connecticut. She earned her PhD from Boston College, where she was introduced to theories of change by Emily Gates. As a researcher and practitioner who studies postsecondary access and success for marginalized students, Kathy has used theories of change to build stakeholder consensus and create collaborative visions for program development and evaluation in US education contexts.

Peter Dahler-Larsen (PhD and Dr Scient. Pol.) is Professor in the Department of Political Science, University of Copenhagen, Denmark, where he teaches evaluation. He is former president of the European Evaluation Society and an honorary member of the Danish Evaluation Society. His international publications include *The Evaluation Society* (2013) and *Quality: From Plato to Performance* (2019). He also edited *A Research Agenda for Evaluation* (2021).

Rick Davies, PhD, is a monitoring and evaluation consultant based in Cambridge, UK. His clients are international development aid organizations. His work includes the management of evaluations, the design of monitoring and evaluation strategies and frameworks, monitoring and evaluation capacity building, and the development of evaluation methodologies.

Thomas Delahais is an evaluator at the worker-owned cooperative company Quadrant Conseil (Paris, France). He evaluates interventions; supports

processes of evaluation, knowledge building, and (re)designing interventions; and trains public agents and project holders in evaluation and design. His areas of work include the pluralism of approaches for evaluating complex interventions, particularly contribution analysis, which he has helped to operationalize; the evaluation of socio-ecological transformation and transition initiatives and their support; and the sociology of evaluation. He is a member of the editorial board of the *Evaluation Journal*.

Heloise Emdon obtained a graduate diploma in public policy and programme evaluation at Carleton University, Ottawa, Canada. As manager of internationally sponsored research projects at Carleton University, she facilitates multidisciplinary teams to develop evidence-informed programs and use monitoring and evaluation to strengthen these programs. She was a student of the 2019–2020 diploma in policy and programme evaluation cohort at Carleton University.

Gordon Freer holds a part-time lecturing position in international relations at the University of the Witwatersrand, Johannesburg, South Africa. He lectured full-time before he was lured away, 20 years ago, to 'practice instead of preach'. In his part-time post, he continues to postulate on his ideals, which are often tempered by his theory-based, real-life evaluations of international development programmes.

Marie Moland Gaarder is the Executive Director of 3ie, leading the organization's efforts to improve the effective use of evidence to inform decision making in international development. Dr. Gaarder has 25 years of experience from international development institutions, including as Manager in the World Bank's IEG, Director of Evaluation at NORAD, and Social Development Specialist at the Inter-American Development Bank. Marie pioneered the initial iteration of evidence gap maps that 3ie subsequently has become known for and has an extensive publication record. Marie is the Co-Chair of the International Development Coordinating Group within the Campbell Collaboration. She holds a PhD in Economics from UCL.

Emily Gates is an assistant professor of evaluation at Boston College. She conducts evaluations in education and public health, publishes theoretical and empirical research on evaluation, and teaches graduate courses in evaluation theory and practice, mixed-methods research, and theory of change. Her research focuses on the implications of systems approaches for evaluation and issues of values, valuing, and equity within the evaluation field.

Eva Grabinski obtained a graduate diploma in public policy and programme evaluation at Carleton University, Ottawa, Canada. Her primary areas of research interest focus on the evaluation of policies, programmes, and services for the provision of social services to vulnerable populations in Canada.

She was a student of the 2020–2021 diploma in policy and programme evaluation cohort at Carleton University.

Hur Hassnain is a senior evaluation advisor for the European Commission, at the Department for International Partnerships' Evaluation Support Service based in Brussels, Belgium. His areas of expertise include the conduct of evaluations in fragile contexts, as well as evaluation of gender, peace, and conflict resolution, rule of law, climate change, renewable energy, agriculture, economic empowerment, education, and protection. He is the founder of the Pakistan Evaluation Association and a board member at the International Development Evaluation Association. He is the lead author of *Evaluation in Contexts of Fragility, Conflict and Violence* published in 2021.

Penny Hawkins is an independent advisor who works with philanthropic, government, and private sector organizations to develop strategic and practical impact measurement systems. Over the past 30 years she has held numerous evaluation leadership roles in a range of different organizations and now serves on several international advisory boards. Penny has a keen interest in developing creative and innovative ways to measure impact that are compatible with uncertainty, emergence, and transformational change.

Tom Hobson is research associate at the Centre for the Study of Existential Risk, University of Cambridge, UK. His work is focused on understanding and mapping the militarization of emerging technologies, particularly biological technologies.

Kurt Houlberg is a Professor of Political Science/Public Policy at VIVE, the Danish Center for Social Science Research. His research covers local government finance, policy and management, local government reform, public policy, public administration and contracting out. His recent publications include articles published in *American Political Science Review*, *European Journal of Political Research*, *Journal of Public Administration Research and Theory*, *Public Administration Review*, *Public Administration*, *Urban Affairs Review*, *and Local Government Studies*.

Andrew Koleros is a Principal Researcher at Mathematica with 20 years of experience in designing and delivering mixed-methods evaluations and programme monitoring, evaluation, and learning systems for small and large social and economic development projects. He brings expertise in using theory-based approaches that integrate complexity concepts and systems thinking into programme and evaluation design processes. His most recent work focuses on providing evaluation and learning services for strategies aimed at catalyzing systems change to address complex societal problems from affordable housing to youth employment to advancing health equity. He has published on his work in peer-reviewed journals and multiple practitioner settings.

Marina Lafer holds a master's degree in public administration from Columbia University's School of International and Public Affairs. She has an undergraduate degree in public administration from Fundação Getulio Vargas in São Paulo, Brazil. She is currently a researcher at FGV EESP CLEAR—the Center for Learning on Evaluation and Results for Brazil and Lusophone Africa. She previously held positions at the Secretary of Planning and Regional Development of the State of São Paulo, Brazil, Fundação Getulio Varga's Project Office (FGV Projetos), and at the Abdul Latif Jameel Poverty Action Lab (J-PAL). She conducted participatory qualitative evaluations for community-based projects in Ecuador, India, and Mozambique.

Frans L. Leeuw is Professor-Emeritus of Law, Public Policy and Social Science Research at the University of Maastricht, Netherlands. He is reviewer and evaluation advisor to the European Union, national governments, and organizations such as the World Bank and IOD PARC. Previously, he was Director of the Netherlands National Institute of Applied Justice and Security Research; Professor of Evaluation Studies, Utrecht University, the Netherlands; Chief-Inspector, Netherlands' Inspectorate for Education; Dean, Humanities and Social Sciences Faculty, Netherlands Open University; Director, Performance Auditing and Evaluation, Netherlands National Audit Office; Associate Professor of Social Policy Research, Leiden University; and Fulbright scholar, University of North Carolina at Chapel Hill. His articles have appeared in journals such as *New Directions for Evaluation*, *German Political Science Quarterly*, *American Journal of Evaluation*, *Evaluation Zeitschrift für Evaluation*, *Utrecht Law Review*, and many others. He is also (co)author of guidance and (hand)books.

Sebastian Lemire, PhD, is an evaluation expert at Abt Associates. He has extensive experience with a broad range of evaluation approaches, quasi- and nonexperimental designs, qualitative and quantitative data collection and analysis methods, as well as systematic evidence reviews. He is an award-winning author of more than 30 peer-reviewed articles and book chapters on a variety of methodological topics, including theory-based evaluation, evaluation capacity building, and mixed-method evidence reviews. He currently serves on the advisory editorial board for *Evaluation* and is a former associate editor of the *American Journal of Evaluation*.

Carlisle J. Levine (PhD) is President and CEO of BLE Solutions, LLC. Levine is an advocacy, peacebuilding, rights, and capacity-building evaluator with more than 20 years of experience working globally and in the United States. In 2011, she launched BLE Solutions, LLC, a boutique evaluation firm, to support the efforts of private and public foundations, international nongovernmental organizations, and governmental and quasi-governmental organizations to strengthen the effectiveness of their practices. Levine has provided

thought leadership on evaluation topics such as using contribution analysis in policy work, measuring champion development, adopting local ownership in evaluation (democratic evaluation), and developing agency-wide measurement systems.

Lycia Lima received a doctorate in public administration and government with emphasis on politics and economics of the public sector from Escola de Administração de Empresas de São Paulo at Fundação Getulio Vargas (FGV/EAESP), with visiting appointments at Massachusetts Institute of Technology (2018–2019) and University College London (2018). She received a master's degree in development economics from the University of London, School of Oriental and African Studies. She has an undergraduate degree in economics from the Universidade Federal de Minas Gerais, Brazil. She is currently a lecturer at FGV/EAESP, Deputy Director at FGV EESP CLEAR—the Center for Learning on Evaluation and Results for Brazil and Lusophone Africa—Associate Researcher at the Center for Applied Microeconomics and Associate Researcher at the Center for Politics and Economics of the Public Sector at FGV. She previously held positions at the World Bank in Washington, DC; the UN in Geneva; the Secretariat of Strategic Affairs of the Presidency of Brazil; and for the government of Minas Gerais State in Brazil. Her research interests are applied microeconomics, particularly public policy evaluation of human capital investment interventions (early childhood, education, and labor market). She also works on research in the areas of urban economics and public finance.

Tom Ling has over 30 years of experience in designing, managing, and delivering complex evaluations focused on innovation, impact, and quality. His clients have included UK government departments and agencies, the European Commission, UNDP, OECD, the World Bank, and many others. He is a senior research leader at RAND Europe and Head of Evaluation. In addition to his current role at RAND Europe, Tom has worked as Head of Evaluation at Save the Children, as a senior research fellow, and at the National Audit Office, and has held various academic posts including Professor Emeritus at Anglia Ruskin University and a Research Associate with Cambridge University. Tom has published widely on evaluation, accountability, and related topics. He co-edited *Performance Audit: Contributing to Accountability in Democratic Government* (2011), following his *Performance Audit Handbook* (2010) and *The Evidence Book* (2017), a critical examination of the use of evidence in public policy and service delivery. With Mary Dixon Wood and others, he is currently co-editing a major book on quality improvement in health care. He publishes regularly in peer reviewed journals. He is currently President of the European Evaluation Society and an advisor to the World Bank's Global Evaluation Initiative.

Igor Lyubashenko is a political scientist working at the Center for Policy Design and Evaluation of the SWPS University of Social Sciences and Humanities (Warsaw). His research interests include conditions that determine the effectiveness of public policies, solutions, and interventions. He is also interested in the theory of complexity and qualitative research methodology (in particular, qualitative comparative analysis and process tracing). He is Principal Investigator in the research project 'Consistency or space for experimentation? Analysis of Poland's policy response to the migration crisis caused by the escalation of Russian aggression against Ukraine', funded by the National Research Center (Poland), and a Project Team Leader in the 'Food Trails' project, funded by the EU Horizon 2020 Programme. At SWPS University, he teaches classes in public management, public policy design, and systemic thinking, as well as methodological courses.

Lara Mani, PhD, is a research associate at the Centre for the Study of Existential Risk, University of Cambridge, UK. Her research seeks to understand the efficacy of various communication methods and strategies for gaining traction for the mitigation and prevention of global catastrophic risks.

Steve Montague is a fellow of the Canadian Evaluation Society, a career practitioner, and an adjunct professor at Carleton University. He specializes in theory-based evaluation and, in particular, studies focused on the reach, engagement, and support of key actors as precursors to outcomes and impacts.

Candice Morkel (PhD) is the Director of the Centre for Learning on Evaluation and Results, Anglophone Africa (CLEAR-AA) and Senior Lecturer in Monitoring and Evaluation at the University of the Witwatersrand, Johannesburg, South Africa. She has 25 years' experience in government, academia, and the non-profit sector, specializing in monitoring and evaluation (M&E) and public policy. As Director of CLEAR-AA, she works with governments and development partners across the African continent to help build better systems of evidence production and use. She has published peer-reviewed articles, book chapters, and edited books on M&E, and supervises local and international master's and PhD students in evaluation. Dr. Morkel has extensive experience in evaluation capacity development, developing curriculum, and teaching in postgraduate and short courses in local and international programmes, including the International Programme on Development Evaluation Training. She has been a board member of the African Evaluation Association, chairperson of the South African M&E Association, and is currently a member of the Board Trustees of the International Evaluation Academy.

Jane Mulcahy is a research fellow with the Research Evidence into Policy Programmes and Practice Project at the University of Limerick and is working on the Greentown programme. She graduated with her PhD in law from

University College, Cork, in 2020 on the topic of 'Connected Corrections and Corrected Connections: Post-Release Supervision of Long Sentence Male Prisoners'. She has worked as a researcher in criminal justice, penal policy, and social justice since 2005. She hosts a podcast called 'Law and Justice' and was awarded the Justice Media Award in the best local radio show/podcast category by the Law Society of Ireland for her three-part documentary series 'Humanising Human Rights' on Ireland's second periodic review under the UN Convention Against Torture in Geneva in 2017. Jane was appointed as a member of the Policing Authority by the Irish Government in July 2023.

Kiruba Murugaiah is a doctoral candidate at Boston College, researching anti-racism in K-12 mathematics teaching and learning. She also explores using decolonial methods in research on teachers in conflict-affected regions in sub-Saharan Africa. In Kiruba's former role as an education technical advisor for a nonprofit operating in humanitarian contexts, theories of change helped her to think about complex problems, design and evaluate programs, and demonstrate value for money to donors.

Catherine Naughton is a research fellow with the Research Evidence into Policy Programmes and Practice Project at the University of Limerick. She is the lead administrator and Principal Investigator – Institutional Review Entity for the Stable Lives Safer Streets Hub (SLSSHub) an All-Ireland policy-led research hub in youth justice. SLSSHub is part of Research Evidence into Policy Programmes and Practice and is a collaboration between UL, Queens University Belfast and the Center for Effective Service. She has worked on the Greentown programme since 2017 and was the lead researcher in the evaluation of the bail supervision scheme for children. The evaluation combined a quasi-experimental design with a theory-driven contribution analysis. She has a PhD in psychology and International Scientific Indexing publications in both qualitative and quantitative methodology.

Mark Oldenbeuving is an independent consultant who specializes in designing and continuously improving projects in complex environments, blending approaches including programme theory development, human-centered design, and evaluation. Mark is a driving force in the design and application of the Actor-Based Change Framework, a new approach for programme theory development in complex settings. This pioneering approach, which was published in the *American Journal of Evaluation*, marks a significant contribution to the field of programme design and evaluation.

Karol Olejniczak, PhD, is the head of the Center for Policy Design and Evaluation at SWPS University and a co-founder of the Polish research company EGO—Evaluation for Government Organizations. His work focuses on programme evaluation, policy design, and using games and behavioral insights in public policy. He was an Ostroms' Policy Analysis Workshop

fellow at Indiana University, Kosciuszko Foundation, and a Fulbright fellow at the George Washington University. He has over 20 years of experience in teaching programme evaluation and conducting applied research for government agencies in the European Union and the US.

Markus Palenberg is the Managing Director of the Institute for Development Strategy. Markus works as researcher, evaluator, and consultant. His research focuses on evaluation methodology such as tools for efficiency analysis, the concept of results chains, results-based management, and causal attribution and contribution. As evaluator, he conducts theory-based evaluations of institutions and their work in the development arena. Markus also consults for programmes and networks on impact strategies, internal governance arrangements and M&E systems. Over the last 15 years, Markus has conducted more than 30 research and consulting assignments in the public and private sector and led more than 10 comprehensive programme and institutional evaluations. From its creation in 2006 to 2009, Markus was Managing Director of the consulting practice of the Global Public Policy Institute (GPPi), a vibrant Berlin-based global governance think tank. Prior to entering the development field in 2005, Markus worked as a corporate manager in one of Germany's largest online marketplaces and, before, as a strategy consultant with McKinsey & Company. Markus' academic background is in the natural sciences, and he worked as postdoctoral researcher at the Massachusetts Institute of Technology for two years after completing his Doctorate in Theoretical Physics in 2000.

Burt Perrin is a thinker, author, and independent consultant with more than 40 years' practical experience assisting governments and other organizations internationally. More recently, he has provided expert advice and quality assurance regarding the design and management of evaluation processes and systems, the design of plans and methodologies, and guidance on interpretation and presentation. He is a recognized leader in the evaluation field internationally, with publications including meaningful approaches to accountability, how to make evaluation useful, and evaluation of innovation. He was Vice President of the International Organization for Cooperation in Evaluation, Secretary General of the European Evaluation Society, and he was awarded its only honorary lifetime membership. He was a founding director and now a fellow of the Canadian Evaluation Society and the American Evaluation Association, where he chaired its awards committee. He serves on the editorial board and is active as a peer reviewer for a variety of international journals, and he holds the credentialed evaluator designation. He likes to take a positive, constructive approach to evaluation, as exemplified in his practice and in various publications and presentations.

Seán Redmond is Adjunct Professor of Youth Justice and is seconded by the Department of Justice (DOJ) to the School of Law, University of Limerick

(UL). He is also a registered social worker and holds a doctorate in governance. He is leading a joint DOJ/UL project to develop the evidence base to improve youth crime policy decision making in Ireland. From 2012 to 2016, he worked in the Research and Evaluation Unit of the Department of Children and Youth Affairs with the lead responsibility for evaluation. He had lead responsibility for the Value for Money and Policy Review of €34 million targeted youth programmes. Over the last 25 years, Seán has held positions as Head of Young Offender Programs at the Irish Youth Justice Service, Assistant Director of Children's Services at Barnardo's, and Director of the PACE Prisoners' Resettlement Service.

Olaf Rieper has an MA in Sociology from the University of Copenhagen and a PhD from the Copenhagen Business School. He is currently an Emeritus Research Director at the Danish Center for Social Science Research. He has conducted evaluations and research on public programmes and organizations for Danish and Norwegian governments and for the EU Commission. The focus has been on welfare services and its organizations as well as on regional development programmes. He has published on research methodology and on the evidence movement. He has been working as a consultant in Ann Arbor, Michigan and in Lyon, France, and was the founding member of the Danish Evaluation Society and the first Chairman of its board.

Patricia Rogers is an independent consultant and researcher, formerly Professor of Public Sector Evaluation at RMIT University (officially the Royal Melbourne Institute of Technology), Australia, and founder of BetterEvaluation.

Mary Tangelder is a focused and passionate researcher, evaluator, and development specialist currently working at the Mastercard Foundation, where she is leading a new interdisciplinary initiative to develop an impact storytelling and narrative practice for the Foundation and their partners. With more than two decades of international humanitarian and development experience, she has led research studies, policy and advocacy initiatives, programme design, and evaluation with leading NGOs, UN agencies, and government bodies. She has lived and worked in more than 30 countries across Africa, the Middle East, Asia, and the South Pacific. She currently resides in Toronto, Canada.

Zazie Tolmer is an evaluation practitioner with 17 years' experience and expertise in theory of change practice. She works with teams from governments, NGOs, foundations, impact investment funds, and mission-driven initiatives to design and implement purposeful learning and evaluation processes and products. She is interested in ways to harness evaluation practice to power, rather than overwhelm, original game-changing efforts.

Peter van der Knaap, PhD, is Director of IOB-Evaluation: the independent evaluation directorate of The Netherlands Ministry of Foreign Affairs. Before, he was Director of SWOV, the National Institute for Scientific Research on Road Safety, and Research Director for the Netherlands Court of Audit. Van der Knaap earned his PhD in administrative science on policy evaluation and decision making within the EU. He is an editorial board member of *Evaluation and Bestuurskunde* and, from 2016 until 2022, chairperson of the Dutch Evaluation Society Vide. Throughout his career, his work and many publications have focused on policy evaluation as a tool for a 'learning government'.

Bob Williams is best known for promoting the use by the evaluation field of approaches from the systems field. However, he has stirred the pot on many other evaluation issues, often from a systemic perspective, including the current addiction to theories of change. In 2014, Bob was awarded the American Evaluation Association's Paul Lazarsfeld Prize for his contributions to the theory and practice of evaluation. His first career was ecological research, but that evolved into an interest in social and environmental movements—and over the past 50 years Bob has managed them, funded them, developed policy for them, researched them, and evaluated them.

Foreword

Emily Gates, Kiruba Murugaiah,
and Kathy Chau Rohn

Our hope for this Foreword is to offer you, the reader, a welcoming invitation to engage with this book in the ways that serve you now, and in the future when new circumstances and questions bring you back. We also encourage you to join the conversation about theories of change that the co-editors and authors in this issue invite us into.

This book covers a lot, by design. The co-editors invited authors from different disciplinary backgrounds and areas of practice to reflect on the strengths, limitations, and future directions of theories of change. Not surprisingly, the essays illustrate a rich variety of experiences, issues, and possibilities for using theories of change within a wide range of contexts. There is no one or best way to engage in a theory of change process or develop diagrams.

In our reading of this book, one big question stands out: WHY are we doing a volume on theories of change?

Together, these essays prompt us individually and as fields to reflect on purposes for theories of change and ask whether these are indeed the right purposes, on what grounds, from whose perspectives, and with what potential consequences. Purposes spark energy and help guide us when making tough decisions about how to develop, illustrate, and use theories of change. What makes for a 'good' theory of change depends, in part, on our purpose.

Purposes also call our attention to different aspects of complexity—a topic and practical challenge evident across these essays. Complexity, in the evaluation literature, tends to be defined or used ontologically to characterize problems, interventions, and systems out there in the world. But complexity also arises epistemologically in relation to our capacity to know and to model change processes, socially and politically as we negotiate different perspectives, and normatively and imaginatively as we envision what futures we desire (Midgley, 1992).

This book showcases multiple legitimate and worthwhile purposes for using theories of change, each with unique aspects of complexity, and brings these to life with practitioners' detailed examples and reflections. We highlight three purposes for theories of change, what attending to complexity means within each purpose, and a few tips from authors.

Purpose 1: To develop a *causal model* of how things work

Since their origins, theories of change have been foundationally about theorizing change. A theory is an explanation of something—in this case, a change process understood as causal processes and mechanisms. This purpose emphasizes logical, critical reasoning, causal relationships, and social science theory and evidence as bases for model development and refinement. Complexity matters in relation to how the world works or how specific interventions or systems operate—what Midgley (1992) calls natural world complexity. Model building is a matter of representing 'what is', typically aiming for truth, accuracy, or validity. Complexity in relation to modeling how things work is a concern when there are multiple interacting components, and the relationships between them are nonlinear and dynamic, such that they produce a change in a way that cannot be fully known or determined (p. 94). Multiple causal pathways exist in nested layers of context, leading some authors to call for systems perspectives (p. 34). A good theory of change would largely depend on the diagram, the theories and evidence that undergird it, how well it captures assumptions, its dynamicity, and other model features.

Some **useful tips** when using theories of change primarily to develop causal models are:

- **Wickedness and root causes of the problem or issue.** Consider what's known, unknown, and uncertain about the underlying problem. Can existing social science theory (p. 15) or root cause analyses inform model development so as not to reinforce change as a 'quick fix' (p. 217)?
- **Context, environment, and nested systems.** Carve out a 'perimeter' for an intervention's contributions within a system of change (p. 15) and develop multiple diagrams with different foci from macro to meso to micro (p. 94).
- **Change mechanisms or processes.** Anticipate 'fuzziness' (p. 129) and 'messiness at multiple levels' (p. 105) when trying to identify detailed understandings of how and why change happens, acknowledging that it may not be possible or desirable to include all the 'finer details around change mechanisms' (p. 129). Rule out alternative explanations (Gaarder).
- **Tracking impact.** Recognize that impact and success may require first a willingness to ask 'tougher questions' (p. 112). Then, they may require answering the questions through a range of approaches such as outcome harvesting, modus operandi, realist evaluation, contribution analysis, and actor-based approaches, among others.
- **Unintended consequences.** Identify and include, whether in narratives, visuals, or supplementary materials, what else might happen, where, and for whom. Several authors point out that these possibilities tend to get left out of visual diagrams.

Purpose 2: To co-develop *shared* model(s) of how we think things work to guide learning and adaptation

This second purpose centers on the process of bringing people together across different perspectives to co-develop models. The primary aim is to develop a 'shared set of assumptions about how things work' (p. 217) with an emphasis on shared. Theories of change provide ways for collective sense making to facilitate learning and guide action. Generating models for this purpose requires careful attention to who is involved, in what ways, and toward what ends. What Midgley (1992) calls subjective world complexity—what any individual thinks, intends, or feels—is at play with the ideal to understand and incorporate this subjectivity. The quality of a theory of change requires looking at the process to see who was included and excluded in the process, how well the process engaged different perspectives and assumptions, and the visual diagram to see how well and usefully it incorporates multiple perspectives in the portrayal of change processes.

Some **useful tips** when using theories of change to develop shared model(s) to guide action are:

- **Consider inclusion and exclusion.** Reflect on whose voices are lifted up or unintentionally suppressed in theory of change creation and use. Failing to consider such considerations can have consequences, especially when working with marginalized or oppressed groups (p. 217), in fragile and volatile contexts (p. 208), or in localized and highly contextualized non-Western settings (p. 129).
- **Consider the evaluator's role and timing.** A range of stakeholder engagement strategies emerged from authors who themselves have different experiences and stances on their role in theory of change processes. Hassnain suggests that evaluator engagement should begin at the design and implementation stages of developing a theory of change, particularly in volatile and fragile environments.
- **Shift power.** Lemire posits that evaluators need to 'flip the script' entirely to create transformative change through equity-based approaches. He argues that increasing stakeholder involvement is insufficient. Instead, evaluators must shift power in the theory of change process to the stakeholders most affected by the programmes themselves.
- **Support collaboration.** Consider various culturally responsive, conflict-sensitive, context-based, and technologically novel approaches to increase collaboration, seek understandings of diverse pathways, intentionally incorporate more voices, and remain attuned to the audiences of the theory of change (p. 129 and p. 208). Experiment with creative collaborative tools such as narrative storytelling (p. 199) and online platforms (p. 225) to incorporate voices across time and location.

- **Surface implicit theories.** Those at the frontlines of change work hold tacit knowledge and latent ideas about why and how something is happening and surface this as a basis for theory development (p. 184), such as when advocates use observations to link advocacy strategies to policy outcomes (p. 25).
- **Creatively express or represent theories of change.** Possibilities are no longer shaped by left-to-right linear depictions of boxes and arrows foundational to logframes and logic models. Technology and virtual online tools for co-construction (p. 225), storytelling (p. 199), and engaging artists (Tangelder) provide multiple ways of visualizing theories of change.
- **Consider not using theories of change.** Several authors (p. 151) mention resistance to or minimal use of theories of change. Williams questions why evaluators rarely use theories of change to model the processes and consequences of evaluations. Listen to participants who would supposedly benefit from a co-development process and consider whether theories of change are cultural and contextual fits or whether there are more appropriate alternatives.
- **Ground process in theories of learning and action.** Given the espoused benefits of theories of change for group learning and action, Williams suggests grounding processes in these theories—including evaluation—to test and refine our own claims about the theory behind using theories of change.

Purpose 3: To co-develop a *shared vision* of how things should be or desired future(s)

The third purpose departs from the above and focuses on 'yearning to imagine new worlds' (p. 184). Theories of change are about what *should be* rather than what is so or why it is so.

Model building here is not about representation but about envisioning, imagining, creating, and dreaming up. As Tangelder notes, 'the theory of change tool was purposely designed to spark imagination about different, better worlds and spotlight what happens when we intervene'. Complexity is most at play in relation to what Midgley (1992) calls 'social world complexity' and 'what ought to be' where the ideal of an inquiry process is 'rightness' in a pragmatic and moral sense. A quality theory of change would not necessarily be about the diagram or the process, but about the catalytic power of the vision to inspire and its moral legitimacy.

A few **useful tips** when using theories of change to envision desired futures are:

- **Imagine new worlds.** Consider beginning with radically open, outside-the-box thinking rather than understanding past or current problems or change processes. Or begin with articulating the vision or goal (p. 184), even if it feels far from the current situation.

- **Explore alternative futures.** Davies, Mani, and Hobson share their use of ParEvo, a 'method of exploring alternative futures or histories using a participatory evolutionary process'. Check it out!
- **Support systems change and transformation.** Changing and transforming large-scale systems require imagining future states and seeing how existing relationships can co-produce emergent outcomes into the future, such as Koleros' work on actor-based systems change.

While these three purposes—causal model building, shared process, and envisioning futures—stood out to us, there are others evident in the essays as well. In reflecting on these purposes, we note how theories of change originated and tend to widely be associated with the first purpose, building causal models. Yet, in these essays, most authors discuss or call for shifts in focus to participation and process, along the lines of the second purpose. And only a few authors take up explicitly normative, imaginative, and future-focused uses of this approach.

As researchers and evaluators, forging ahead regardless of whether the purpose fits the circumstances or insisting on some right or best way of doing theories of change sets us up for risks. These essays prompted us to consider how getting clear on the purposes of theories of change could help us attune to the aspects of complexity that matter most and consider the quality of our work. As you read, we hope that you feel the same level of excitement and energy as we did when considering where theories of change might go in the future.

Reference

Midgley, G. (1992). Pluralism and the legitimation of systems science. *Systems Practice*, *5*(2), 147–172. https://doi.org/10.1007/BF01059938

Part 1

Introduction

*Andrew Koleros, Marie-Hélène Adrien
and Tony Tyrrell*

'Nothing as practical as a good theory'

Kurt Lewin (1943)

A hallmark of humanity is our desire to try to make sense of the world around us. From the earliest questions we pose as children to the foundations of scientific inquiry, humans are constantly searching for explanations about how and why things happen as they do. Within different disciplines—from philosophy to economics to political science, among others—various approaches have been designed to make sense of how and why change occurs. The field of evaluation developed its own perspectives for understanding individual and societal change processes and for explaining how particular projects, policies, or programmes (that is, interventions) can lead to observed change.

The predominant approach developed by the evaluation profession in pursuit of this endeavour is commonly referred to as *theories of change*. Since the first published articles on the use of theory to inform evaluation planning more than 50 years ago, evaluators have used theories of change and theory-based approaches to articulate the logic underpinning how a particular intervention is intended to bring about a desired change as a core component of evaluation design. Over time, evaluators began using theories of change for other important purposes beyond causal analysis, such as for developing a shared understanding of the vision of an intervention and the key strategies designed to achieve it.

As stakeholders outside of the evaluation community began to see the usefulness of this approach, theories of change were adopted and adapted more widely, with uses expanding from its origins in evaluation design to include programme design and management, communications, routine monitoring, and adaptive management. The wider use and application of theories of change for multiple purposes has thus led to substantial diversity—both inside and outside of the evaluation community—in the different ways in which theories of change are developed, represented, and used.

In some ways, the flexibility of theories of change and the ability to adapt them for multiple purposes are some of their main benefits and help to explain why they have been so widely adopted. At the same time, this diversity has resulted

in a good deal of ambiguity and little consensus around how best to define, develop, represent, and use theories of change. Indeed, numerous reviews have documented this general lack of agreement, particularly within the international development community (Stein & Valters, 2012).

In recent years, a growing number of programme theorists have begun to address this gap through more explicit discussions around the essential components and characteristics of theories of change (Davies, 2012, 2018; Mayne, 2018; UK Department for International Development, 2012). Notably, Funnell and Rogers's seminal work *Purposeful Program Theory* provides a comprehensive history and the contemporary landscape of the different approaches to theory of change development and use, suggesting guidance and offering examples for different types of interventions. Since its publication in 2011, it has become a leading resource—if not *the* resource—for those interested in better understanding the approach.

Given the robustness and relevance of this literature, the intention of this volume is not to reinvent the wheel. In fact, there is no wheel to reinvent; although there is growing agreement on the core components and characteristics of theories of change, the literature also suggests that there is, *in reality*, no one way to develop and use theories of change. The usefulness of the approach is inherent in how it is used.

As theories of change continue to be increasingly and more widely used within a growing number of disciplines, our book aims to re-examine this approach that has been a foundation stone of Western evaluation for the past few decades from multiple perspectives. To ensure the book included diverse perspectives, ranging from experienced evaluators to those involved in programme design, implementation, monitoring, and evaluation—and to privilege the voices of newer and emerging practitioners—we asked a larger number of potential contributors to provide a short essay (as opposed to the more traditional format of 10 to 12 full-length chapters) responding to these three questions:

- How are you using theories of change in your work?
- What limits to their use have you found?
- What further adaptation does the approach need to ensure that it remains relevant and useful in the future?

Overall, 33 women and men responded to our call and submitted a total of 26 essays, including contributions from Global North and Global South practitioners, and from many newer and emerging practitioners, complemented by more experienced evaluators and experts on the topic. We initially envisioned that each author would provide an essay sharing their own experiences using theories of change responding to one or more of the questions above. However, authors not only wrote about their own personal experiences using theories of change, but also provided broader and often more critical reflections on how

the approach is currently defined and used across a broader landscape of policy, programme, and evaluation contexts.

The way in which these various contributors responded to our call thus widened the focus of the book. It moved the content beyond a narrower perspective on the strengths and limitations of theories of change among a subset of practitioners—who potentially have a more positive view of theories of change and regularly use them in their own practice—to a broader perspective on the favourable and potentially harmful ways that the wider practitioner community thinks about and uses theories of change. We revisit and expand on these insights in our concluding chapter.

By sharing these diverse perspectives through this wider lens, the book thus aims to provide theory of change practitioners—from evaluators to programme theorists, planners, and implementers—with critical insights on the relevance and applicability of theories of change across different contexts through concrete examples of their uses, as well as insights to a wider audience outside the evaluation community on how best to use or adapt the approach.

Structure and content

The authors in this edition used multiple terms to describe the approach in their respective essays, perhaps reflective of the varied ways in which the approach is understood in practice. The concept of 'theory-based evaluation' may have started with the work of the German-English sociologist Karl Mannheim, who in 1934 coined the concept of 'principia media', that is, time and space restricted assumptions about what makes planning effective (Mannheim, 1951). In the 1950s, cognitive psychologists began discussing 'lay theories of behaviour', which unpacked the beliefs and causes people attribute to behaviour. Around this time, Leeuw described how the Popperian tradition of critical rationalism, which seeks to make theories more testable, also influenced the development of theory-based evaluation (Leeuw, 2003). Finally, during the 1960s policy scientists and evaluators started to think in terms of models, theories, and assumptions underlying policies and programmes.

Building on this tradition, the first reference to the theory of change approach is often attributed to a publication that emerged from a 1995 roundtable event at the Aspen Institute, where Carol Weiss coined the term 'theory of change' to evaluate complex community change initiatives (Kubisch et al., 1995). With wider adoption over time terms such as 'logic models', 'outcome chains', 'programme theory', and 'theory-based evaluation' have now become part of the theories of change lexicon. Funnell and Rogers (2011) discuss the 'theory of change' and the 'theory of action' as two components of 'programme theory' that have now been adopted by many programme theorists as the preferred nomenclature for the approach, while the term 'theory of change' remains more common in some sectors, such as international development (Funnell & Rogers,

2011; Freer & Lemire, 2019). More recently, the terms 'causal pathways' or 'causal impact pathways' are often included as a key component of a theory of change (Koleros & Mayne, 2018; Stame, 2004).

For consistency and ease of reading, the editors have opted to adopt the terms theory of change or theories of change to refer to the approach throughout this edition. While theory of change was most commonly used by authors—and is arguably the term most commonly used in the literature—we have pluralised the term where appropriate to reflect that a singular theory often is not sufficient to capture the complexity of many programmes and the multiple purposes for which the approach is used. As is further explored throughout many of the essays in this book, a particular programme might have one theory of change that is used for programme design purposes, another that serves more of an external communications purpose, while another informs monitoring, evaluation, and learning exercises. Mayne (2018) shared a similar viewpoint: 'a key concept for an intervention is that… there can usefully be several depictions of the ToC [theory of change], typically with different amounts of detail' (p. 2). In this edition, we contend that theories of change may better reflect the current practice and have opted for this term throughout the book; however, at times we also use 'theory of change' where more grammatically appropriate.

The different ways that authors in this edition defined theories of change within their essays provide insights into how the wider practitioner community understands the approach. To visually depict the diversity in definitions, we developed the word cloud presented in Figure 0.1, where the words that appear larger are those that were mentioned more often across the multiple definitions.

Authors defined theory of change both in terms of what it is (that is, what it looks like and how it is depicted) and by its use. In describing what it is, they differentiated between the theory itself and the overall 'theory of change approach'.

Figure 0.1 Word cloud of the multiple author definitions of 'theories of change'

They defined the theory itself using terms such as models, constructs, descriptions, articulations, and stories. A theory of change approach was generally described as a process to explicitly identify how programmes are supposed to work to address a specific problem, setting out the different possible sequences of steps from interventions to a desired set of outcomes situated within wider systems in which these programmes operate. In defining the approach by use, authors mostly referred to it as an approach, a tool, or an instrument to 'open the black box of a programme' and obtain a common understanding of the change pathways underpinning a programme and the underlying programme mechanisms, as well as the contextual conditions and assumptions needed to fulfil the cause-effect relationships along those pathways.

Use of the approach was defined not only in terms of how it is currently used both inside and outside the evaluation community, but also in terms of how it could or should be used and cut across multiple dimensions. We identified five core themes across these dimensions that provide the overarching structure of the book.

The first two themes concern the use of theories of change within an evaluation context:

- Theme 1: Considerations in using theories of change to establish causality
- Theme 2: Using theories of change to understand wider societal change processes

The next two themes reflect on the use of theories of change beyond the field of evaluation:

- Theme 3: Adapting theories of change for use in broader public policy contexts
- Theme 4: Applying theories of change approaches for multiple purposes

The last theme cuts across the others and discusses the role of stakeholders within the process:

- Theme 5: Developing theories of change that reflect multiple stakeholder perspectives

Summaries of the main conclusions according to these five themes are further described in the remainder of this section.

Theme 1: Considerations in using theories of change to establish causality

The five essays under this theme explore the use of theories of change to establish causality with reference to a wide range of current factors and concerns, including: the need to put change (rather than the intervention) at the centre of the evaluation process; the limitations of the theory-based approach and

tendencies towards oversimplification given increased complexity in the broad, political economy; the need to adopt a systems-led approach to allow for the interconnectedness of things; and the need to socialise theories of change to enhance their relevance and increase the possibility of stakeholder buy-in.

Thomas Delahais supports the use of social science theories to isolate an intervention's possible contributions to change within a system. Having isolated possible contributions to change, various strategies can be applied to better elaborate and test causality. In adopting this contribution analysis approach, Delahais purports to place change (rather than the evaluated intervention) at the centre of the evaluation process. **Carlisle Levine** suggests that the contribution analysis approach championed by Delahais can support evaluators to develop realistic narratives about how change happens. She asserts that contribution analysis, when underpinned by robust theories of change and by an inclusive approach, is a tool that offers evaluators the opportunity to develop 'detailed and credible stories' about the change process.

In his essay, **Markus Palenberg** presents theories of change as qualitative conceptual tools that support the deconstruction of cause-and-effect relationships within a system, help explain their interrelationships, and explain underlying assumptions. Following this process, theories of change can help in understanding, systematically analysing, and making predictions about cause-and-effect relationships. Palenberg also notes the limitations associated with how comprehensively any model can capture real-world phenomena and cautions that this must be borne in mind when the evaluator interprets what it is that theories can reveal about real-world cause-and-effect relationships. His caution about the limitations of theories of change in comprehensively describing the real world is echoed in **Peter Dahler-Larsen's** essay, where he notes that, despite the role played by critical thinking and logical sequencing in the design of theories of change so that they can best represent reality, logic alone does not rule reality. Consequently, the practice of evaluation takes place in the space between controlled, logical thinking on the one hand and social practices on the other. He suggests that improved dialogue and socialisation of the theory-making process can further enhance the potential of theory-based evaluation. **Tom Ling** draws on ontological insights to identify practical ways to improve evaluations using theories of change. He also advocates for the adoption of key principles in support of better-quality evaluation in which theories are best applied: engaging stakeholders, incorporating systems thinking, iterating the theory of change, prioritising theory building, and being humble and curious.

Theme 2: Using theories of change to understand wider societal change processes

Moving beyond a core use of theories of change in establishing causality, five essays discuss how theories of change are being used to evaluate expanded scopes

and broader contexts beyond traditional programmes, particularly within societal change processes. Over time, evaluators' understanding of individual and societal change processes has evolved in response to emerging theories from the fields of behavioural science, complexity science, and systems thinking, among others. As evaluators' understanding of 'how change happens' has shifted over time, so too has their use of theories of change to support their evaluative efforts. At the same time, evaluators have increasingly expanded the scope of their evaluative activities, moving from evaluations of single projects or programmes to evaluations of multiple programmes within wider social and ecological contexts, on topics from human rights to climate change and the Sustainable Development Goals. As evaluators have expanded the scope of their evaluative efforts into these new domains, this has resulted in a wide diversity of practice in theory of change development and use within the evaluation community.

Patricia Rogers writes about the use of theories of change particularly for interventions in complex systems and the challenges that presents for developing, representing, and using theories of change. Her essay explains that interventions in complex systems operate under conditions of ongoing uncertainty. She then posits that a different approach is needed to use theories of change in these situations. Such an approach would effectively support ongoing single-loop and double-loop learning and adaptation and assist in managing the features of complex systems rather than being used for planning tightly prescribed implementation of 'what works'. **Andrew Koleros** provides a concrete example of some of these principles in practice in the philanthropic sector, making the case that an actor-based approach to theories of change can assist foundations in designing strategies that focus on achieving systems change. He argues that evaluators can play a role in this process, but it will require some shifts in how they typically engage with foundations: engaging earlier in the strategy design process, adopting a design mindset as well as evaluation mindset, and embracing the uncertainty of systems change efforts.

Frans Leeuw writes about the use of theories of change within the fields of artificial intelligence (AI) and big data (BD). He describes that although evaluating 'black boxes' is familiar territory for evaluators, very rarely has a realist, theory-driven approach to evaluation been applied when BD and/or AI are involved in policy making and implementation. Leeuw outlines a potential approach to unpacking AI black boxes through a theory of change approach. **Penny Hawkins** and **Zazie Tolmer** write about their work using theories of change in the context of impact investing. They describe how adapting the approach to a market-based setting forced them to move beyond a conventional programmatic approach. They argue that the approach they adopted was strategically useful in articulating, testing, improving, and communicating their approach to impact in a market context and helped investors think about how to scale. **Jens Andersson** describes the use of theories of change for sustainable business endeavours, arguing that the approach can be an important tool to support these

processes, but that it must be adapted to a private sector context. He explains that this requires pragmatic use of the tool and a learning mindset, while accepting that it may come at the cost of loss of rigour in the way theories of change have conventionally been conceived and used in a development context.

Theme 3: Adapting theories of change for use in broader public policy contexts

Five essays engage with the adaptation of theories of change and their use in broader public policy settings. Whereas theories of change were originally applied to the evaluation of bounded interventions (projects, programmes) to help explain how such interventions were expected to lead to specific development change, the boundaries of the field of play to which theories of change are now applied have expanded. A key theme reflected in the essays is that the use of theory-based approaches to evaluation in broader public policy and administration contexts can result in overly simplified models. The authors present various suggestions through which this tendency can be mitigated, such that a theory-based approach can add value.

Candice Morkel discusses the need for a systems-oriented approach in the context of increased complexity in public sector planning, with particular reference to the utility of current practice in the use of theories of change in macro public sector planning in Africa. While welcoming the spirit of these initiatives, Morkel cautions against the adoption of weak, linear models that fail to adequately embrace the complexity of political, cultural, and socioeconomic contexts in post-colonial Africa. She advocates for the adoption of a more locally grounded, indigenised systems perspective that avoids oversimplification. Similar issues are observed, with a continental shift in focus, by **Lycia Lima and Marina Lafer**, reflecting on the use of theories of change to plan and manage national multiyear development planning in Brazil. They find that the use of theories of change in a context characterised by multiple, interactive interventions, has tended towards oversimplification, suggesting there may be merit in using nested theories of change that have been developed through inclusive, participatory processes. They caution, however, that even this approach will likely encounter hurdles if it is to be meaningfully applied in the context of complex, multiyear development strategies.

Peter van der Knaap widens the lens to reflect on the relevance of theories of change in the context of ever-increasing emphasis on sustainability in development. He argues that the adoption of the UN Sustainable Development Goals introduces further complexity into the development arena, providing an even greater challenge for theory-based evaluation. To meet this challenge, van der Knaap advocates for a system-based theory of change approach that goes beyond 'measure' and 'effect' to also specify the necessary system requirements for sustainable impact. The analytical lens is pulled into a more localised context

by **Kurt Houlberg and Olaf Rieper** to explore the use of theories of change in assessing the efficacy of structural reform in local government in Denmark. They identify two possible ways in which a theory-based approach can be more directly applied to future, complex reform processes: favouring a systemic route to better capture what they refer to as 'the multi-purpose, multi-stakeholder, multifaceted, and multi-actor implementing nature' of the local reform effort; as well as using a nested approach through which tailored theories of change would be designed for specific aspects of the reform. Finally, **Karol Olejniczak and Igor Lyubashenko** elaborate findings from their empirical study into the potential use of simple heuristics to support citizen engagement in drafting theories of change. Their essay elaborates the results of a case study into the experimental application of an inclusive approach to drafting theories of change for the food waste problem in the Warsaw metropolis. The study finds that the 'policy lab' innovation, which links the evaluation practice with applied behavioural science, represents an effective tool for citizen engagement in designing change, for generating ideas and for handling complexity.

Theme 4: Applying theories of change approaches for multiple purposes

Beyond these specific uses around public policy, stakeholders outside of the evaluation community have increasingly adopted theories of change for different purposes along the policy and/or programme cycle. These include planning and design, monitoring and adaptive management, and organisational learning. The wider use and application of theories of change for multiple purposes has led to even more diversity in the ways that theories of change are developed and represented.

Mary Tangelder discusses how the approach can be used as a design tool to help funders and their grantees shift from short-term, discrete activities to addressing deeper, systems-level change. Drawing from her personal experience, she builds a case for why we need to rethink and re-imagine how to use theories of change to inspire action, build shared ownership, drive learning and adaptation, and ultimately accelerate the path to change. Moving from programme design to programme management, **Mark Oldenbeuving** discusses the ways in which theories of change can serve as an effective tool for programme delivery, and how at present they are often not used to their full potential in this area. He describes ways in which theories of change can be more actively used in programme management to improve programme delivery, providing concrete examples from his own work. In her essay, **Marie Gaarder** describes the use of a theory of change as an approach to drive organisational learning within the International Initiative for Impact Evaluation (3ie). From this framing, she describes the underlying theory of change of 3ie and—with the benefit of hindsight—describes how well it has been fulfilled over the past 14 years. She goes

on to reflect on what 3ie and other specialised institutions in the evidence space can do moving forward to further realise their collective theory of change.

In their contributions, **Burt Perrin** and **Gordon Freer** take a wider view on this topic and reflect on how the use of theories of change has evolved over time, leading in some instances to their misuse. Although theories of change often are promoted as something needed by every programme and evaluation, in all possible situations, Perrin explores whether there is really a universal need in all cases. Using an operatic analogy, Perrin describes that, while theories of change in many situations can be very valuable, this is not always the case; in practice, few are actually used or are helpful, and in some cases they can be detrimental. He identifies some key considerations when contemplating use of theories of change so that they add value. On a similar note, Freer uses the hammer-and-nail analogy of the law of the instrument to explore the minimal utilisation of theories of change within programmes despite their almost mandatory inclusion in programme proposals, designs, and evaluations. He considers reasons for this lack of use and explores potential solutions, contending that theories of change can be used in programme design as complementary programme management tools, adding value to reporting, adaptation, and learning, as well as contributing more broadly to refining social science theories.

Finally, **Bob Williams** shines light on the evaluation community itself, writing about the use of theories of change (or lack thereof) to guide evaluation use. He describes how theories of change, rightly or wrongly, are virtually a compulsory part of contemporary evaluation design alongside a focus on evaluation usefulness, outcomes, and consequences. He goes on to investigate why—if theories of change are as useful as the evaluation community claims—do so few evaluators develop one to demonstrate or focus their claims for the usefulness or consequences of their evaluations? With a critical lens on the evaluation community and his own practice as an evaluator, Williams speculates on possible reasons for this state of affairs—some good, some bad, and some arguably unprofessional.

Theme 5: Developing theories of change that reflect multiple stakeholder perspectives

Theories of change can help programme managers and evaluators make sense of how change happens, explain how human interventions have caused or have contributed to these changes, and posit what this means for the future. However, such interpretation is inherently biased by the evaluator's mental constructs, values, beliefs, and identity, potentially distorting both the interpretation of data collected and ensuing findings. To mitigate such biases, and making use of both theory and practical examples, five essays advocate for engaging stakeholders in designing theories of change, proposing different methods for such engagement and capturing the multiple benefits—and pitfalls—of co-designing. These five essays offer an interesting perspective on the challenges in co-designing theories

of change, including the need for additional resources—namely time and financial resources—to bring stakeholders together and the need for an effective moderator with a strong understanding of group dynamics, as well as conflict resolution and a sensitivity to cultural and gender differences in communication.

In his essay, **Sebastian Lemire** argues that co-designing theories of change is a useful approach towards promoting equity in the evaluation process, facilitating a shared understanding of how a programme is intended to work. Lemire argues that co-designing facilitates shifting the power dynamics between the evaluator and the evaluand, and reduces disparity, exclusion, and discrimination while allowing the voices of beneficiaries to be heard. Echoing Lemire's conclusions, **Jane Mulcahy, Catherine Naughton, and Sean Redmond** contend that co-designing theories of change can lead to programme improvement, using an example from their work with the Intensive Family Programme pillar of an intervention for children involved in adult crime networks in Ireland. The co-design of theories of change by programme managers and programme beneficiaries brought to bear distinct stakeholders' visions and aims for the programme, resulting in rich discussions and the redesign of the programme components. For Mulcahy, Naughton, and Redmond, the co-designing process offers stakeholders a safe space where programme assumptions can be assessed and re-engineered in real time.

Hur Hassnain's essay reaffirms the benefits of co-designing theories of change, particularly in fragile contexts. Hassnain discusses why and how evaluations in contexts of fragility, conflict, and violence must aim to reconstruct theories of change together with all the stakeholders to ensure that the evaluation remains fully grounded in the country's rapidly changing context. Similarly, **Steve Montague, Heloise Emdon, and Eva Grabinski** argue that co-designing theories of change can increase conceptual clarity about evaluation focus and context. They share their experiences of virtually co-designing a theory of change during the COVID-19 pandemic through several rounds of e-workshops. Their experience sheds light on the benefits of co-design for strategic reflection and programme and evaluation design improvement. Noting the limits of the more traditional visualisation of theories of change, **Rick Davies, Tom Hobson, and Lara Mani** argue in their essay that narratives can also be a useful vehicle for describing different theories about the future. They describe how ParEvo.org enabled the participatory construction of a branching structure of alternative narratives about biological research risks and the opportunity to accumulate lessons from other similar exercises in the future. The essay highlights how the collaborative approach offered the opportunity to accumulate, compare, and learn from the experience of developing these theories of change.

User's guide

Given the diversity of perspectives shared in this book, the way in which we grouped essays by these five themes was an editorial decision which emerged

thematically as we engaged with the content of each essay. But it goes without saying that each essay provides its own insights beyond its thematic grouping and no doubt there will have been many other ways we could have grouped this interesting collection of essays. As such, a reader might choose to dip in and out of the essays within these thematic groupings based on their interests and curiosity. We encourage readers to dip in and out frequently over time as their own interest in and experiences with using theories of change evolve.

References

Davies, R. (2012). Criteria for assessing the evaluability of theories of change. Retrieved from http://mandenews.blogspot.com/2012/04/criteria-for-assessing-evaluablity-of.html

Davies, R. (2018). Representing theories of change: A technical challenge with evaluation consequences. CEDIL Inception Paper 15. Retrieved from https://mande.co.uk/wp-content/uploads/2018/09/2018-08-31-Inception-Paper-No-15-MandE-NEWS-PDF-copy-2.pdf

Freer, G., & Lemire, S. (2019). Can't see the wood for the logframe: Integrating logframes and theories of change in development evaluation. *Canadian Journal of Program Evaluation*, *33*, 3.

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory*. Jossey-Bass.

Koleros, A., & Mayne, J. (2018). Using actor-based theories of change to conduct robust contribution analysis in complex settings. *Canadian Journal of Program Evaluation*, *33*(3), 292–315.

Kubisch, A. C., Weiss, C. H., Schorr, L. B., & Connell, J. P. (1995). Introduction. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 1–22). Aspen Institute.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, 24, 5–20.

Lewin, Kurt. (1943) Psychology and the Process of Group Living. *Journal of Social Psychology*, 17, 113–131.

Mannheim, K. (1951). *Man and society in an age of reconstruction*. Martinus Nijhoff.

Mayne, J. (2018). Developing useful theories of change: Evergreen briefing note. Retrieved from www.researchgate.net/publication/323868372_Developing_and_Using_Useful_ToCs

Stame, N. (2004). Theory-based evaluation and types of complexity. *Evaluation*, *10*(1), 58–76. https://doi.org/10.1177/1356389004043135

Stein, D., & Valters, C. (2012). Understanding 'theory of change' in international development: A review of existing knowledge. The Asia Foundation and the Justice and Security Research Programme. Retrieved from www.theoryofchange.org/wp-content/uploads/toco_library/pdf/UNDERSTANDINGTHEORYOFChangeSteinValtersPN.pdf

UK Department for International Development. (2012). Broadening the range of designs and methods for impact evaluations. DFID Working Papers, no. 38. Retrieved from www.oecd.org/derec/50399683.pdf.

Part 2

Considerations in using theories of change to establish causality

1 Setting the stage for contribution claims

Thomas Delahais

Introduction

How do you evaluate impact when the evaluated intervention is a plausible component of expected change, but is unlikely to be a major driver—say, the role of research in major socioeconomic developments? This often happens in complex situations, where change depends on many interacting drivers and overlapping interventions. Theory-based evaluations (TBE) tend to focus on the intervention and keep change in the background, as a *context* or a *final impact*. They may overstate attribution of impact to the evaluated intervention by maintaining out of their scope the many external reasons that may explain change. An alternative approach is to go the other way round and put *context* centre stage and consider how an intervention can influence it. The first section below illustrates how we came to this 'inversion' in the context of evaluating the impact of research. Making such a move brings new challenges, though: change is complex, and human beings cannot fully embrace the multitude of drivers and interventions at stake in change and how they interact. So, how is it possible to draw assumptions about contributions to change if change cannot be well understood? The second section explores the use of social science theories to carve out a 'perimeter' for an intervention's contributions within complex change, one in which it will be easier to make sense of the drivers at stake and plausible contributions to them.

The contribution of research towards sustainable forest management

A few years ago, we evaluated the role of two research centres, the Center for International Forestry Research (CIFOR) and the French Agricultural Research Centre for International Development (CIRAD), in the expansion of sustainable forest management practices in the Congo Basin in the last 20 years (Delahais et al., 2014; Delahais & Toulemonde, 2017). It was clear from the start to all stakeholders involved in this evaluation that research was not the main driver of observed change in forestry practices, but they all had convincing claims about its role nonetheless.

Our approach to this evaluation was contribution analysis (Delahais & Toulemonde, 2012; Mayne, 2000, 2012). In the 2010s, contribution analysis

'came of age' and helped evaluators to consider complexity more seriously by introducing or perfecting some key concepts that proved useful for the evaluation in question. First, change (for example, in forestry exploitation) happens continuously because of powerful drivers and many overlapping interventions. Second, *desired* change, if observed (for example, more sustainable forest management), is unlikely to be due to an intervention alone. Third, an intervention's contribution can never be certain, if only because pathways to change are manifold and uncertain, and their ability to trigger change depends on configurations of stakeholders, contexts, etc. The humble role of evaluators is then to 'increase the confidence' (Befani & Mayne, 2014) that an intervention has a role (or has not) towards these changes and explain how and why.

Contribution analysis brought the notion of 'causal package' (Mayne, 2012) to deal with the multiplicity of causes to observed change. An intervention may not be sufficient or necessary to trigger a change, but a package of drivers, among which the intervention sits, can. However useful, this concept was still more adapted to situations where the intervention is likely to be a key driver. Our case was different: several other large-scale interventions were aiming at similar changes and were better candidates than research to explain change *prima facie*.

The solution was similar to that proposed by Funnell and Rogers (2011) of making a distinction between a *theory of change*—this term referring here to 'the central mechanism by which change comes about for individuals, groups, and communities' (p. 31) and a *theory of action* that connects the intervention to this 'overall system' at different steps (Figure 1.1).

In this case, the approach we used meant:

> delineating the set of major trends that supposedly drove the observed changes in forest management practices, [starting] with the stakeholders whose practices were ultimately expected to change (timber companies, entrepreneurs of the informal sector) and [moving] upstream towards those most likely to have influenced these practices (client companies, Congo Basin country governments… NGOs, international institutions, bilateral donors, and international public opinion. (Delahais & Toulemonde, 2017, p. 372)

To assess the role of research in these trends, we iteratively developed plausible statements about contribution to change—or contribution claims (CCs).

In contribution analysis, important efforts are devoted in the initial phase of theory building to collect and flesh out claims about if, how, and why an intervention can lead to changes. Over the course of the evaluation, the plausibility of the claims is progressively substantiated. Only those claims that are plausible or important enough to the stakeholders are tested (theory-testing phase; Ton, 2021). Defining 'plausible' therefore is important. A common definition is that

Classical TBE

CHANGE is CONTEXT

Conditions & Risks   C&R   C&R

Intervention → Outputs → Outcomes → Impact
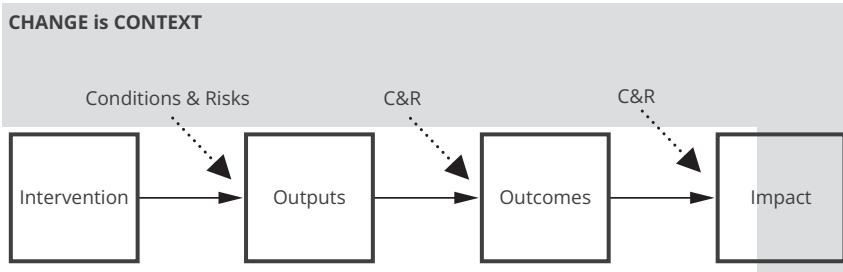
Contribution analysis with causal packages

CHANGE feeds impact

| Set of drivers | C&R | Set of drivers | C&R | Set of drivers | C&R | Set of drivers |
| Intervention | | Outputs | | Out-comes | | Impact |
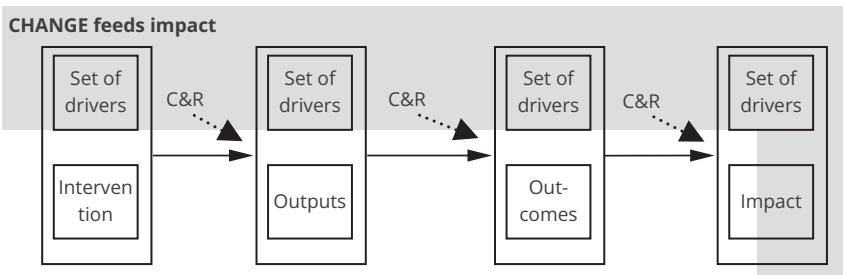
Contribution analysis with change on centre stage

Intervention

Many drivers

Many other interventions

Towards desired or undesired change

System of change (follows its own trends)
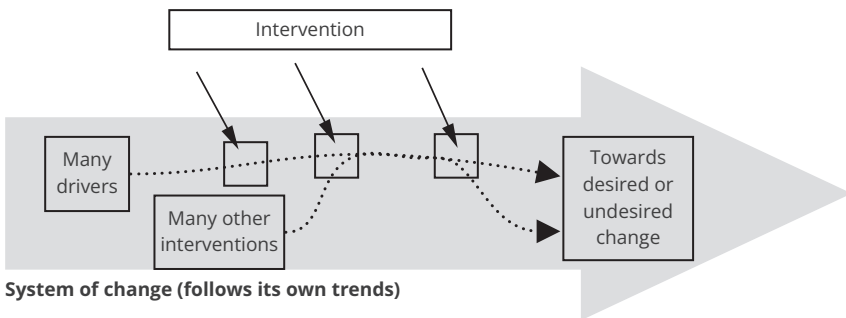
Figure 1.1 Distinction between theory of change and theory of action

of Connell and Kubisch (1998), in which a claim is plausible if it is 'supported by prior empirical evidence or common sense'. The construction of an overall theory of change allowed us to expand this definition and consider that a contribution is plausible when the intervention is possibly connected to the major drivers and trends that constitute change. In fact, this evaluation changed how we would define CCs, not only as plausible statements but as the set of plausible pathways connecting change and the intervention, *starting from change* (Box 1.1).

Box 1.1 Contribution to the development of certified wood products

Ecolabels for wood products have their roots in the concept of sustainable forest management, which CIFOR helped frame and advocated for in the 1990s as an alternative to strict conservation strategies ('framing' claim). They rely on certification frameworks which are calibrated on local data. CIRAD was the only research centre able to provide the necessary input for the Congo Basin region ('data' claim). For their part, timber companies were willing to develop new industrial processes to tap this new market. They funded pilot projects led by CIFOR and CIRAD to do so—with success ('solutions' claim).

Framing change and contribution through different theories

The above CCs are presented as 'bridges' located at the interface between change and the evaluated intervention. Starting from change allows for more plausible claims, but it also comes with new challenges—specifically having a much better command of change, not only of the intervention. This is not easy; change is complex. How it unfolds is the result of multiple interactions of many drivers within *systems of change*. It is not possible to entirely make sense of it or of the directions it is following or the full set of causal factors leading to it (Morin, 1988; Preiser, 2016). Palenberg (Essay 3) also describes some of the difficulties he faced in developing a theory of change with a 'system perspective', as well as the benefits of such an approach.

In the case of CIFOR and CIRAD, change was largely described according to the stakeholders' view that international organisations and states were prominent drivers of change. This was a sensible choice, given the time and budget. But in hindsight, this may have kept us blind to other perspectives on change, to which CIFOR and CIRAD could have contributed (for example, community-led change). And it may have led us to overstate the contribution of research to change. As Ling rightfully points out in his essay, a theory of change should not be confused with the real world, but it should help structure our understanding of change. The question then is: How can complex change be framed in a way that will allow for the development of plausible claims about contribution?

Science uses theories and concepts to understand and explain its objects (Gingras, 2017). In evaluation, theories can similarly be used to frame an evaluand. Leeuw and Donaldson (2015) have clarified how theories can be 'knitted' into evaluation work, namely by 'integrating parts of (at first sight non-related or loosely coupled) theories'. They argue, however, that 'not all the types of

theory… can be knitted together; only can those that are more or less similar, in terms of their type of content and orientation' (p. 474). This suggests that the result should be consistent; but from a complex perspective, there is no such thing as a general change theory because systems cannot be reduced to only one dimension (Morin, 1988). Different, competing, or even antagonistic patterns of change are not exclusive and may cohabit within the same system (Morin, 2013).

Theories do not always need to fit together. Rather, they act as multiple windows onto complex change. Evaluators can use their explanations and properties to understand the changes to which interventions are expected to contribute, and how and why interventions fit into these changes. Based on the above, there appear to be three major ways in which theories can be used when developing CCs: (1) to 'carve out a perimeter' for CCs, within a system of change, in which contribution is more likely and easier to comprehend than outside (Figure 1.2); (2) to define or clarify the pathways—and the associated conditions—through which an intervention can affect this portion of the system of change; and (3) to develop empirical tests to assess contributions and their significance.

There are different strategies of using theories to develop CCs. In line with theory knitting, the first one would be to look for **congruence** to increase plausibility, that is, 'where, empirically, … Theory A is valid and where Theory B is valid [and] mechanisms that can explain why this is the situation' (Leeuw & Donaldson, 2015, p. 474). Sectoral theories can set the stage for CCs by describing change and explaining if, how, and why an interaction between the intervention and the system of change can happen. Policy theories will help to
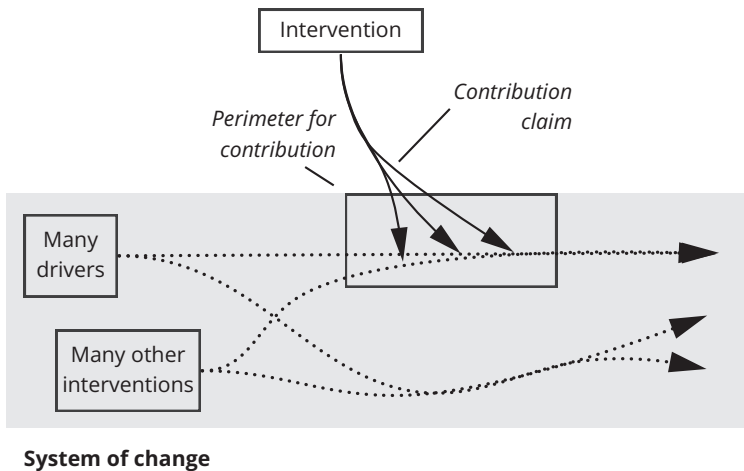

Figure 1.2 Contribution claims within a system of change

understand how certain instruments or mechanisms are expected to affect society; behavioural theories will help to understand how people will react to these instruments, and so on. Embedding CCs into these different types of theories helps qualify their plausibility. The process of rejecting theories or contextualising when they are relevant or not will also help detail not only if, but when, CCs are plausible.

Alternatives, especially in participatory settings, include **scaffolding**—using concepts to start thinking about the perimeter in which contributions are plausible but without any expectation to keep them in the final product (as one would remove a scaffold once a building is finished). Meta-theories, like the COM-B behavioural theory (Michie et al., 2011), work well with this strategy because they are generic or mainstream enough to take a first look at a situation or to trigger exchanges with stakeholders. Confrontational theories can be used through **antagonistic dialogue** to reveal assumptions about what is expected from the intervention or clarifying language that obfuscates what is expected from an intervention (Cornwall, 2010).

In this phase of theory building, theories should not be used to totally validate or discard claims (this is the role of theory testing), but only to increase or decrease the level of plausibility associated with them. Indeed, evaluators are limited in the repertoire of theories they use (they are not omniscient, they are limited in time and budget, and they have their own values which limit what they are willing to rely on). Besides, seemingly relevant theories are identified based on an observation of at least some changes (or the anticipation of observing them). However, this observation may be quickly obsolete in a context of emergent change—all pathways to change cannot be known and the constant interactions between the intervention and the system of change may lead to unexpected changes.

An example: The evaluation of MOPAN

The Multilateral Organisation Performance Assessment Network (MOPAN) assesses the organisational performance arrangements of multilateral organisations (MOs) such as the United Nations Agencies on behalf of its member countries. In 2019, MOPAN issued a logic model that sequentially links its assessments to improved performance of MOs and, ultimately, to the achievement of the UN Sustainable Development Goals. This chain of events was unlikely, at least on the MO side: MOs have their internal and external dynamics of change, and there was limited evidence that they were using the MOPAN assessments at all. The initial plausibility of any claim related to MOPAN's contribution to increased performance and, as a consequence, to MO's effectiveness in the field was very low. Yet, significant efforts had been made in recent years to try to make MOPAN assessments more useful to MOs, and it was therefore important to investigate this question.

In the evaluation (Quadrant Conseil, 2022), a system of change in the performance of MOs was outlined, and an area was carved in which an impact of donor-led assessments was plausible. This initial perimeter was defined using a mix of **congruent** international relations theories and concepts (accountability and strategies of influence) and theories of policy making (new public management as an instrument of reform of MOs through results-based management; Geri, 2001). MOPAN is presented as an instrument of 'accountability', but the meaning of that word in this context was unclear. The comparison between two **antagonistic** visions of accountability 'as a virtue' and 'as a mechanism' (Bovens, 2010) was used to clarify that MOPAN was initially established with a view of accountability 'as a virtue' (the mere existence of assessments ensures that MOs will seriously consider performance arrangements in the spirit of new public management), but expectations had changed such that assessments were now to act 'as a mechanism' (they should be used to influence MOs into reinforcing their arrangements). Studying the evolution of accountability relationships highlighted that accountability-to-donors was now only one mechanism of accountability among others.

The second step was to understand how this mechanism could influence change within MOs. Evaluation theory was used for **scaffolding**: could instrumental, conceptual, symbolic use be expected from the report or the process (Alkin & King, 2017)? MOPAN reports are not evaluations, though. Their methodology is rooted in the institutionalist view that change happens when formal arrangements are made (for example, MOs do more evaluations once they have an evaluation policy). For this reason, they focus on these formal arrangements, but listing those does not provide much information to the assessed MOs. This limited any instrumental use of assessments to organisations that did not already have such arrangements in place, did not know whether they had such arrangements, or had not recently reviewed them. We used **antagonistic** theories related to the influence of indicators (positivist versus constructivist) to rule out any possible impact in terms of 'international benchmarking' between MOs, where the publication of indicators leads to public response and a commitment to improve (for example, Radaelli, 2020), but also to reveal how influential a process of collaborative indicator building could be (as MOPAN had done on the prevention of sexual exploitation and abuse).

Still, preliminary investigations had shown some plausible cases of use, which needed to be explained. Several policy theories were used, some of which were only partially compatible (for example, political entrepreneurs in stakeholder coalitions and epistemic communities, capacity building, or knowledge brokering), to clarify how and where in the decision-making processes MOPAN could be used. For instance, assessments could be used as ammunition in internal debates by parties of reform (such as stakeholders advocating for a more ambitious evaluation policy).

These theories also led to a third level of explanation centred on the agency of individuals in the use of knowledge. But it was acknowledged that this was not feasible for the evaluation to investigate at that fine-grained level; this part therefore remained a stub.

At the end of this theory-building process, a series of claims related to contribution was issued, each supported by one or several pathways, which were assessed against their plausibility and possible significance. Theories were among the sources used to develop the empirical tests through which the use of MOPAN assessments was appraised in a series of MOs.

Conclusion

This essay has argued that CCs are like bridges between a system of change and an intervention. Evaluators are typically required to have a good understanding of the interventions, but if they are to make plausible claims, they need to have a good command of change, too. Theories can be used as windows into the system of change but also as ways to better qualify the plausibility of CCs in complex settings. Doing so before in-depth testing is crucial to ensure that the evaluation will be relevant and efficient in the use of its means.

Theories do not replace the assumptions of stakeholders, which are foundational in developing a theory of change, as discussed by Levine (Essay 2): they support, complement, compete with, or challenge them. They act as an additional layer of understanding and as a way to engage with stakeholders to obtain some level of deliberation, as emphasised by Dahler-Larsen (Essay 4). In the case of MOPAN, institutionalism was frequently discussed with member countries and this helped to shed a different light on MOPAN's internal methodological debates. But they are also a safeguard to avoid having the evaluator or another stakeholder impose their framing, voluntarily or not. Indeed, evaluators can be unaware that their view of what is plausible and what is not is embedded into an often-implicit theory or system of values (Schwandt, 2019). Using multiple, divergent theories can also help to collectively acknowledge that any claim can only be plausible within a certain view of how change unfolds. This emphasis on transparency ultimately contributes to the credibility of the evaluation. In the words of Edgar Morin (1988), 'Theory should account for what makes possible the production of theory itself' (p. 10).

References

Alkin, M. C., & King, J. A. (2017). Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. *American Journal of Evaluation*, *38*(3), 434-450. https://doi.org/10.1177/1098214017717015

Befani, B., & Mayne, J. (2014). Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation. *IDS Bulletin*, *45*(6), 17–36. https://doi.org/10.1111/1759-5436.12110

Bovens, M. (2010). Two concepts of accountability: Accountability as a virtue and as a mechanism. *West European Politics*, *33*(5), 946–967. https://doi.org/10.1080/01402382.2010.486119

Connell, J. P., & Kubisch, A. C. (1998). *Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects, and problems*. The Aspen Institute.

Cornwall, A. (2010). Introductory overview—Buzzwords and fuzzwords: Deconstructing development discourse. In A. Cornwall & D. Eade (Eds.), *Deconstructing development discourse: Buzzwords and fuzzwords* (pp. 1–18). Practical Action Publishing, Oxfam.

Delahais, T., Flichy, A., & Ekoumou, C. (2014). *Evaluation of CIFOR's and CIRAD's contribution to sustainable forest management in the Congo Basin*. Center for International Forestry Research, Indonesia.

Delahais, T., & Toulemonde, J. (2012). Applying contribution analysis: Lessons from five years of practice. *Evaluation*, *18*(3), 281–293. https://doi.org/10.1177/1356389012450810

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

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. Jossey-Bass.

Geri, L. R. (2001). New public management and the reform of international organizations. *International Review of Administrative Sciences*, *67*(3), 445–460. https://doi.org/10.1177/0020852301673004

Gingras, Y. (2017). Qu'est-ce qu'une science? In M. Silberstein (Ed.), *Qu'est-ce que la science… Pour vous* (pp. 119–124). Éditions matériologiques.

Leeuw, F. L., & Donaldson, S. (2015). Theory in evaluation: Reducing confusion and encouraging debate. *Evaluation*, *21*(4), 467–480.

Mayne, J. (2000). Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation*, *16*(1), 1–24.

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, *18*(3), 270–280.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(1), 1.

Morin, E. (1988). Le défi de la complexité. *Chimères*, *5*(1), 1–18. https://doi.org/10.3406/chime.1988.1060

Morin, E. (2013). *La nature de la nature (La méthode, tome 1)*. Seuil.

Preiser, R. (2016). *Critical complexity: Collected essays*. De Gruyter. https://doi.org/10.1515/9781501502590

Quadrant Conseil. (2022). *Evaluation of the Multilateral Organisation Performance Assessment Network (MOPAN)*. Retrieved from www.mopanonline.org/aboutus/evaluations/#:~:text=MOPAN%20underwent%20an%20independent%20external,supporting%20an%20effective%20multilateral%20system.

Radaelli, C. M. (2020). Regulatory indicators in the European Union and the Organization for Economic Cooperation and Development: Performance assessment, organizational processes, and learning. *Public Policy and Administration*, *35*(3), 227–246. https://doi.org/10.1177/0952076718758369

Schwandt, T. A. (2019). Post-normal evaluation? *Evaluation*, *25*(3), 317–329. https://doi
.org/10.1177/1356389019855501

Ton, G. (2021). Development policy and impact evaluation: Learning and accountability
in private sector development. In H. Zafarullah and A. S. Huque (Eds.), *Handbook of
development policy*. Edgar Handbooks in Development.

2 Using theories of change to assess causality in a policy change context

Carlisle J. Levine

Introduction

It is difficult to assess how policy change comes about. Establishing causality in policy change requires considering all actors and factors involved, the time required to bring about policy change, and policy change's frequently nonlinear path. This is echoed in Delahais' essay, which examines the role of theories of change in evaluating interventions in complex settings. Contribution analysis—a theory-based impact evaluation approach developed by John Mayne and discussed in Delahais' essay—is one approach that addresses these challenges. Using it, evaluators develop detailed and credible stories about how a change came about, taking into account all relevant perspectives. Underlying these stories is the evaluators' meticulous work with advocates and policymakers to develop well-triangulated theories of change, identifying different actors and factors' unique contributions and carefully testing causal links and assumptions. This essay provides a brief overview of contribution analysis's use in establishing causality in a policy change context and then focuses on contribution analysis's grounding in robust theories of change, delving into actions required to ensure their rigour in advocacy contexts.

Using contribution analysis to assess causality in complex contexts

In complex contexts, it is very difficult to assess causality by proving that initiative A caused result B—a variable approach to causality assessment. The evolving nature of these initiatives does not allow for holding variables constant. Identifying credible control or comparison groups in complex contexts is potentially possible, when making causal linkages between discrete activities and short-term results. However, it quickly becomes quite difficult to accomplish when assessing the contribution of broader initiatives to long-term policy change.

Process approaches to assessing causality, in contrast, examine how a change unfolded (Maxwell, 2004). As also noted by Palenberg (Essay 3), they allow evaluators to carefully examine the logic and rationale underlying theories of change and test them against alternative hypotheses. In that way, evaluators can

assess the credibility of causal claims, identify the role a particular initiative played in bringing about a desired outcome, and compare that contribution to the contribution of other initiatives and events. Contribution analysis is one of these process approaches.

John Mayne developed contribution analysis for use in public sector performance management (Wimbush et al., 2012). Since then, evaluators have used it to assess causality in complex contexts, such as those related to research use, knowledge exchange, and advocacy (Delahais & Toulemonde, 2012; Kane et al., 2017, 2021; Koleros & Mayne, 2019; Lemire et al., 2012; Mayne, 2019; Stocks-Rankin, 2014; Wimbush et al., 2012).

To undertake contribution analysis, Mayne laid out six steps. Building on Mayne's original model (2012), Kane et al. (2021) modified the steps for a policy change context and added a seventh step that was originally suggested by Wimbush et al. (2012):

1. Describe the specific causal questions to address.
2. Develop robust theories of change and possible alternative explanations.
3. Gather evidence on the theories of change and possible alternative explanations.
4. Assemble the contribution story and assess the contribution claim, as well as challenges to it.
5. Seek out additional evidence.
6. Revise and strengthen the contribution story.
7. Use findings to learn and make improvements.

More information about using contribution analysis to assess advocacy's impact can be found in Kane et al. (2017, 2021), of which I am a co-author. In these writings, we wrestled with contribution analysis's second step: *Develop robust theories of change and possible alternative explanations*. In our experience, advocates have implicit theories of change that guide their work, but they rarely articulate them. We also know that, given their work's fast pace and the brief policy windows during which they can make a difference, they are unlikely to offer evaluators enough time to elicit an explicit, robust theory of change from them. Yet, we recognise the fundamental role a theory of change plays in process approaches to assessing causality, with its pathways connecting actions to shorter- and longer-term results, its causal linkages explaining those connections, and its assumptions articulating what must be true for the linkages to hold. To resolve this dilemma, in our practices, we use alternative approaches to uncover and test the causal linkages and assumptions in advocates' implicit theories of change.

In this essay, I describe approaches for eliciting implicit theories of change and testing causal relationships. I offer guidance for ensuring the robustness of a theory of change. I draw on lessons learned from our earlier experiences using

contribution analysis to examine how state and national policy change came about in the United States, as well as from more recent experiences focused on local and national policy change in low- and middle-income countries.

Eliciting implicit theories of change

To assess causal linkages, our evaluation teams first needed to discover ways to elicit advocates' implicit theories of change, which we would then test. Advocates' theories of change would provide what Palenberg refers to in his essay as 'intervention perspectives', describing how they believed their efforts contributed to their desired policy changes. However, we needed to elicit these theories of change without using too much of the advocates' time. Document reviews, observing advocacy strategy meetings, workshops, focus group discussions, and interviews all proved to be useful approaches, each offering strengths and weaknesses that are described below.

We began by *reviewing relevant documents*, such as advocacy materials, advocacy strategy meeting notes, news pieces, reports, and grant-related documents, where relevant. In one case, a participant had recorded meeting minutes and was willing to share them with us. These minutes gave us a view into how the advocates perceived the challenges they faced, the rationale behind the actions they selected to address them, and the differences they believed their actions made. The minutes also offered a record of the evolving policy context and how the advocates adjusted their strategies in response to that context. By following the minutes, we were able to develop timelines for different advocacy initiatives that showed correlations between advocate actions and desired results—a variation on results chains. These correlations pointed to hypotheses about causal relationships that we could then test using other data sources to help us build a robust and explicit theory of change. In this case, by reviewing documents, we developed the following timeline summarised in Figure 2.1 (see Kane et al., 2017).

The timeline suggested a results chain containing advocates' actions and policy responses. We used it to identify causal linkages, assumptions, and related questions to test them. How did the US congressional offices decide that introducing this bill was important and something on which they wanted to lead? How did they develop language for the bill? Did they have assistance from outside entities? What sparked the US Agency for International Development's interest in the bill? What encouraged the US senators to introduce a companion bill in the Senate? We used interviews and workshops to answer the above questions, carefully triangulating data sources, noting that interviewee interests could lead to biases in whom they reported as influential in achieving these policy responses. Using content analysis, we assessed correlations between the advocates' draft bill language and the language that emerged in the bill.
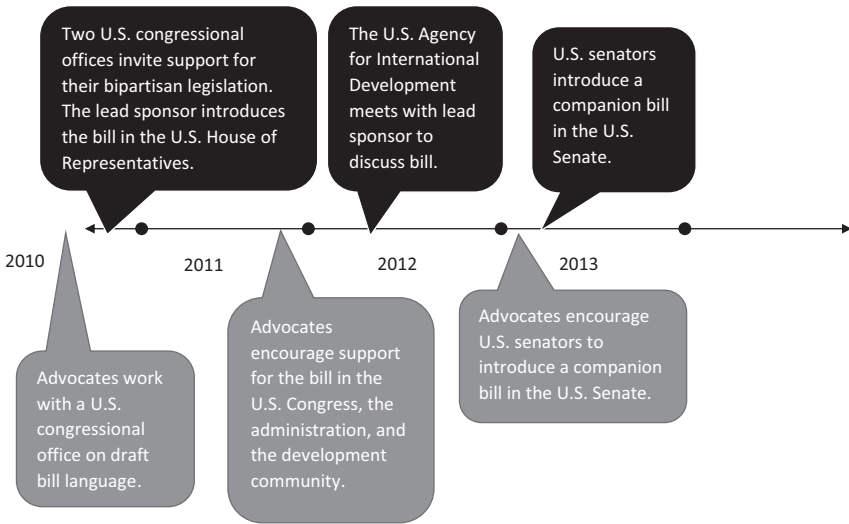
Two U.S. congressional offices invite support for their bipartisan legislation. The lead sponsor introduces the bill in the U.S. House of Representatives.

The U.S. Agency for International Development meets with lead sponsor to discuss bill.

U.S. senators introduce a companion bill in the U.S. Senate.

2010

2011

2012

2013

Advocates work with a U.S. congressional office on draft bill language.

Advocates encourage support for the bill in the U.S. Congress, the administration, and the development community.

Advocates encourage U.S. senators to introduce a companion bill in the U.S. Senate.

Figure 2.1 Advocacy and policy change timeline for a piece of US national legislation

While having access to meeting minutes in this evaluation proved valuable to developing our initial draft timeline and identifying the questions we wanted to address, meeting minutes are not always available to evaluators. In other cases, we have relied on other relevant documents—advocacy materials, reports, and news pieces—to construct timelines. Where these have been less detailed, we have relied more heavily on subsequent interviews and other data collection methods to fill in the gaps. We recognised that *attending regularly scheduled advocacy strategy meetings* could give evaluators additional insights into how advocates believe change will come about, without demanding additional time from the advocates. As echoed in other essays in this volume, doing this successfully required developing trust-based relationships with advocates, so that advocates could feel comfortable letting an external evaluator hear them discuss their strategies and tactics. This is a common approach used by internal evaluators who accompany advocacy initiatives. It points to the possibility of internal evaluators—with their close relationships with advocates—and external evaluators—with their outsiders' eye—partnering on this type of evaluative work.

When advocates' time has allowed, we have *conducted group or individual interviews* with them to elicit their theories of change. Through group interviews, evaluators can help advocates articulate a consensus theory of change. However, depending on power dynamics within a group and louder and softer voices, some perspectives might get overlooked, if evaluators do not facilitate the discussion carefully. Individual interviews can be easier to schedule and

ensure that every voice is heard, although receiving input that may diverge from the dominant narrative regarding how change is coming about may still rely on evaluators' abilities to establish trust with an advocate. With information gathered from interviews, evaluators will need to note areas of consensus and divergence and identify ways to test both.

Testing causal relationships

To test emerging hypotheses about causal relationships, we conducted workshops with advocates and their allies. These workshops, when well attended (often resulting from encouragement provided by those leading the advocacy initiative), can be an example of evaluation as intervention: they offer another venue for advocates and their allies to reflect on their work and consider paths forward, thus offering value to both evaluators and advocates.

During these workshops, we posted timelines emerging from our document review, such as the one shared above, along with questions about causal relationships between timeline activities. As advocates and their allies reviewed these, they talked with each other about how changes have come about. This helped us add causal linkages and assumptions to the timeline results chains.

By triangulating our data sources via interviews with policymakers and others working on the same policy issue, we further tested the causal linkages that we discovered through our document reviews, workshops, and other data-gathering opportunities. These interviewees have often pointed to additional relevant documents for our review.

In the case described above, we sought out interviews with US congressional staff members involved in advancing the legislation of interest, as well as relevant US Agency for International Development staff members. Because they had constructive relationships with the advocates working to advance this legislation, they were willing to talk with us and described, from their perspectives, the various roles different actors played.

To gather alternative explanations for how a policy change has come about, our line of questioning for these interviewees began with the policy change of interest, probing backwards through the results chain the interviewee described to understand how they perceived change to have happened. Palenberg (Essay 3) describes this as 'designing theories of change from a "systems perspective"', in which 'the intervention has—*a priori*—no special status among the contributing causes' and through which 'all relevant causes that contribute to a specific change should be described'.

Through this process, we elicited new elements to add to our emerging theory of change. We also noted if and how an interviewee described the importance of our advocacy initiative of interest, thus increasing our understanding of how that advocacy initiative's contribution to a policy change fit with the contributions of others.

Table 2.1 Methods for eliciting and testing theories of change

Method	Benefits	Considerations
Document review	There are no demands on anyone's time, beyond providing documents. Documents present perspectives regarding how a change came about.	Evaluators must understand whose perspective a document represents. Documents' claims require testing.
Observation of advocacy strategy meetings	There are no additional demands on advocates' time. Evaluators get to hear how advocates view policy change challenges and how they will address them.	Evaluators must earn the trust of advocates to be invited to these meetings and to have advocates speak openly, while they are there. Partnering with an internal evaluator who has advocates' trust may be useful. Because this method is not intended to interfere with meeting proceedings, if evaluators have questions during the meeting, they need to find other opportunities to raise them.
Group interviews with advocates	Evaluators can help advocates articulate a consensus theory of change.	Group interviews can be difficult to schedule, since they require finding a time convenient for all participants. Good attendance may depend on encouragement from an advocacy initiative's lead. This may lead to participants emphasising the initiative's effectiveness more than they would have otherwise. Evaluators need to listen for this. Depending on power dynamics within a group and louder and softer voices, some perspectives might get overlooked, if evaluators do not facilitate the discussion carefully. Evaluators must understand interviewees' perspectives and interests and interpret the information they provide accordingly. Evaluators need to note areas of consensus and divergence and identify ways to test both.

(*Continued*)

Table 2.1 (Continued)

Method	Benefits	Considerations
Individual interviews with advocates, allies, and policymakers	Individual interviews can be easier to schedule. In individual interviews, it is easier to ensure that every voice is heard. Interviews can be used to triangulate data sources and further test causal linkages and assumptions. Interviewees might recommend additional documents for review.	Receiving input that may diverge from the dominant narrative regarding how change is coming about may still rely on evaluators' abilities to establish trust with an advocate. Evaluators must understand interviewees' perspectives and interests and interpret the information they provide accordingly. Evaluators will need to note areas of consensus and divergence and identify ways to test both.
Workshops with advocates and their allies	These workshops, when well attended, can be an example of evaluation as intervention: they offer another venue for advocates and their allies to reflect on their work and consider paths forward, thus providing value for both evaluators and advocates. Workshop conversations can help explore causal linkages and assumptions related to how change has come about. Workshops can bring together different perspectives to come up with an agreed theory of change among advocates and their allies.	Workshops can be difficult to schedule, since they require finding a time convenient for all participants. Good attendance may depend on encouragement from an advocacy initiative's lead. This may lead to participants emphasising the initiative's effectiveness more than they would have otherwise. Evaluators need to listen for this. Depending on power dynamics within a group and louder and softer voices, some perspectives might get overlooked, if evaluators do not facilitate the discussion carefully. Evaluators will need to note areas of consensus and divergence and identify ways to test both.

A key question we include near the end of interviews like these asks: 'Absent [the advocacy initiative of interest], what, if anything, would have been different about [the policy outcome of interest]'? This question helps to clarify the advocacy initiative's contribution, if any. In our experience, responding to a question like this, interviewees can describe advocacy initiatives as contributing significantly

to important steps in a policy change process; helping a change come about faster; serving as one more piece of a much vaster effort to bring about change—not critical, but equally important among other advocacy efforts; or unimportant to bringing about a desired change. With each response, we consider the biases of the respondent and then analyse all responses together to make an overall assessment.

Table 2.1 summarises methods for collecting data to help develop and test theories of change, their benefits, and considerations for their usage.

Ensuring the robustness of a theory of change

To successfully elicit an explicit and robust theory of change that can inform a contribution analysis process, evaluators can use document reviews, observation, interviews, and workshops. However, these data-collection methods rely heavily on people's perceptions, which can call into question a theory of change's completeness and accuracy. Each respondent will have their own perspectives and personal biases and be prone to recall bias. Advocates may be more likely to identify the importance of their own initiatives, while policymakers may be unlikely to admit that they have been influenced by others.

A few approaches can help evaluators overcome these challenges. To mitigate their own biases and those of others involved in the evaluation, evaluators must thoroughly and carefully triangulate findings by drawing on different data collectors, sources, and methods, as well as different data analysers and analysis approaches. As part of this, evaluators must understand their data sources' interests and interpret the information they provide accordingly. Having insider knowledge about a policy change, the policymaking process, and the actors involved will help evaluators correctly interpret what they are hearing. Evaluators may have this knowledge or can partner with internal evaluators or others who do. Alternatively, holding sense-making sessions to review data with people knowledgeable about an issue and context may also help evaluators arrive at more accurate interpretations.

All these approaches require time, resources, experience, expertise, and—particularly—experience and expertise in the context in which the evaluation is taking place. The greater the investment in these elements, the more likely evaluators will emerge with a robust and credible theory of change that accurately describes how a policy change came about and the contribution of a specific advocacy effort to it.

References

Delahais, T., & Toulemonde, J. (2012). Applying contribution analysis: Lessons from five years of practice. *Evaluation*, *18*(3), 281–293. https://doi.org/10.1177/1356389012450810

Kane, R., Levine, C., Orians, C., & Reinelt, C. (2017). *Contribution analysis in policy work: Assessing advocacy's influence*. Center for Evaluation Innovation. www.evaluationinnovation.org/wp-content/uploads/2017/11/Contribution-Analysis_0.pdf

Kane, R., Levine, C., Orians, C., & Reinelt, C. (2021). Contribution analysis: A promising method for assessing advocacy's impact. *New Directions in Evaluation*, *2021*(171), 45–57. https://doi.org/10.1002/ev.20471

Koleros, A., & Mayne, J. (2019). Using actor-based theories of change to conduct robust evaluation in complex settings. *Canadian Journal of Program Evaluation/La Revue canadienne d'évaluation de programme*, *33*(3), 292–315. http://doi.org/10.3138/cjpe.52946

Lemire, S. T., Nielsen, S. B., & Dybdal, L. (2012). Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations. *Evaluation*, *18*(3), 294–309.

Maxwell, J.A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3–11.

Mayne, J. (2012). Contribution analysis: Coming of age? *Evaluation*, *18*(3), 270–280. https://doi.org/10.1177/1356389012451663

Mayne, J. (2019). Revisiting contribution analysis. *Canadian Journal of Program Evaluation/La Revue canadienne d'évaluation de programme*, *34*(2), 171–191. https://doi.org/10.3138/cjpe.68004

Stocks-Rankin, C. R. (2014). *Reflective literature review of contribution analysis*. Institute for Research and Innovation in Social Services.

Wimbush, E., Montague, S., & Mulherin, T. (2012). Application of contribution analysis to outcome planning and impact evaluation. *Evaluation*, *18*(3), 310–329.

3 My perspective on theories of change

Markus Palenberg

Introduction

This essay lays out my personal understanding of theories of change and how I use them in practice. It highlights one aspect I consider important to render this a truly value-adding instrument to the evaluator's toolbox: the *perspective* taken when designing theories of change.

Before beginning, I must admit that I do not think the term *change* in the name *theories of change* is entirely appropriate. To me, it suggests that there must always be some form of change, which excludes the important possibility that a system may react by preserving its status quo—for example, when I pedal my bike and steer to avoid falling over. I will, however, continue using the term for lack of a simple alternative.

The essay is structured as follows. In the following section, I sketch my understanding of theories of change and introduce an example. I then describe two perspectives that can be taken when designing theories of change: the intervention and the system perspective. In the final section I argue in favour of the system perspective.

How I understand theories of change

My way of thinking about—and making use of—theories of change is shaped by my education and research in physics, where models are used all the time. To me, theories of change are *qualitative conceptual tools* that can help *understanding*, systematically *analysing*, and *making predictions* about *cause-and-effect relationships* in the real world.

Theories of change do this by breaking down the cause-and-effect relationships of interest into smaller parts. They describe how these parts are causally related—to each other and to their context—and they explain what assumptions the model is based on.

Theories of change—like all models—have limitations in how well they can describe real-world phenomena. Even under the best of circumstances, a theory of change will only describe specific aspects of the real world. And it will only be able to do that for specific purposes. This limitation sets important boundaries

for the validity of any specific theory of change. In consequence, theories of change can only add value if their limitations are remembered when interpreting what they can tell us about the real world.

In more specific terms, I consider a theory of change to represent a simple *causal model* of the *reaction* of a *system* to *external influences*. Of course, these terms deserve further explanation:

- The s*ystem* consists of selected parts and aspects of the real world that are of interest.
- With *model*, I refer to a systematic conceptual description that attempts to approximate the system at hand. Theories of change are *causal models* because they attempt to describe and explain cause-and-effect relationships within the system and—to some extent—also between the system and its surroundings.
- *Influences* are everything that affects the system. *External influences* originate from outside the system. These can be activities purposefully implemented with specific objectives in mind, such as the activities planned and implemented in projects, programmes, and institutional reform processes. External influences can also be any other activity, changing condition, or evolving situation originating from beyond the system boundary that affects the system in some way. In contrast, *internal influences* originate from within the system and describe causal connections between different parts of the system.
- Finally, in a theory of change, the *reactions* of the system to external influences are approximated by how the explicitly modelled states of affairs describing the system are affected (that is, how they change or remain stable under external influences).

Admittedly, this is still a rather generic and unspecific characterization of theories of change—partly on purpose. It reflects the flexibility I believe is required to make theories of change useful tools, so that they can best serve specific purposes at hand. The way I understand theory of change also draws on more rigorous and structured approaches, foremost on the causal calculus introduced by Judea Pearl (Pearl, 2009; Pearl & Mackenzie, 2018) but also on functional equation modelling and system dynamics.

As in those approaches, I usually summarize important elements of a theory of change by means of a diagram, as shown in examples below. I consider this an intuitive way to visually highlight the most important causal connections in the theory of change. In addition, I write up the theory of change as a narrative, referring to the diagram for illustration. When done, the diagram and the narrative should fully explain all the above-mentioned elements of the theory of change.

In my work to date, I have not made much use of theories of change prepared by others (that is, programme managers) prior to the commencement of
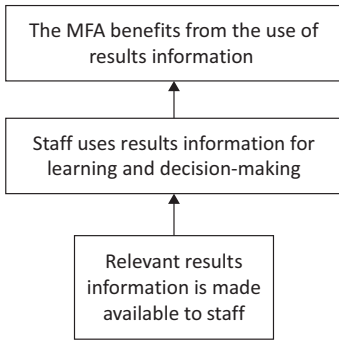
```
┌─────────────────────────────────┐
│  The MFA benefits from the use of │
│       results information         │
└─────────────────────────────────┘
               ↑
┌─────────────────────────────────┐
│  Staff uses results information for │
│    learning and decision-making   │
└─────────────────────────────────┘
               ↑
      ┌──────────────────┐
      │  Relevant results │
      │ information is made│
      │  available to staff│
      └──────────────────┘
```

Figure 3.1 Initial theory of change

evaluations in which I have been involved. This may seem surprising because most of my evaluations are theory based. However, in most instances, there simply was no theory of change. In the few instances when a theory of change had already been developed, we (the evaluation team) usually did not find it particularly useful in relation to our evaluative information needs; hence, we ended up adapting it significantly or simply developed a new one, as explained in the following case.

A theory of change for learning from results

In 2019, I led an evaluation for the Ministry for Foreign Affairs of Finland (Palenberg et al., 2019). In this evaluation, the team was asked to assess how information about the results of Finnish development policy and cooperation had been used in the past and what could be done to strengthen evidence-informed learning and decision making in the future. There was no explicit theory of change. However, implicitly, the initial idea was that *results information*—the umbrella term we introduced for evidence of *what* results had been achieved and *how* they had been achieved—was what caused learning and decision making that, in turn, was thought to contribute to increased performance of the ministry towards its mandate (Figure 3.1).

However, once we understood the topic better it became evident that this was not the most important causal pathway. After a first round of interviews and literature review, we designed a theory of change around the basic causal relationships shown in Figure 3.2.[1] At its heart, this theory of change used an adapted version of the COM-B framework introduced by Michie et al. (2011).

1 This is a simplified version of the theory of change used in Palenberg et al. (2019).
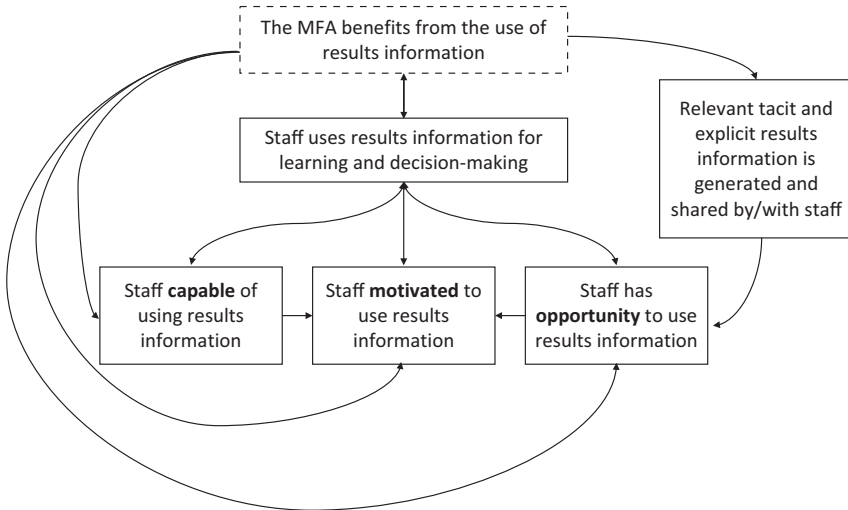
Figure 3.2 Theory of change including basic causal relationships

John Mayne, the evaluation team member leading the theory of change design, had also used this framework in other work (Mayne, 2019). It explains behaviour change—the 'B' in COM-B—by means of three groups of conditions: capability, opportunity, and motivation (COM), defined as follows (Michie et al., 2011):

- *Capability* is the individual's psychological and physical capacity to engage in the activity concerned. It includes having the necessary knowledge and skills.
- *Motivation* is all those brain processes that energize and direct behaviour, not just goals and conscious decision making. It includes habitual processes, emotional responding, as well as analytical decision making.
- *Opportunity* is all the factors that lie outside the individual that make the behaviour possible or prompt it.

These three conditions for greater use of results information for learning and decision making were adopted as essential by the evaluation team. The initial starting point—the availability of results information—was reduced to one factor among several, feeding into the opportunity staff had for making use of it.

The full theory of change we used had more detail than what is shown in Figure 3.2. In addition to the diagram, it explained what we meant by what was in the boxes and what assumptions were behind the arrows connecting them. To

avoid overloading the diagram, we decided not to add detail on how all activities influence any of the boxes or the connections in the diagram (as in the case of the lowest box in Figure 3.1); instead, we listed them separately as 'supporting activities'. We also decided that the causal step from use of results information (the middle box) to increased organizational performance (the top box) represented a hypothesis we would not be able to verify empirically (hence the dotted border).

We used this theory of change as the primary framework throughout the evaluation, adapting it repeatedly in light of what we learned. It guided our interviews and desk review, and it was used to structure our findings, synthesis, and sense making.

I bring up this example because it reflects two different perspectives that can be used when designing theories of change, as explained below.

Perspectives

When designing theories of change, an obvious starting point is the specific 'intervention' at hand. This reflects typical evaluation questions, such as How effective was the intervention in contributing to intended results? or What was the impact of the intervention?

In such cases, the term *intervention* refers to a set of planned or already implemented activities that are part of a project, programme, or some other change process. During theory of change design one can then start with the intervention and explain how its activities cause subsequent effects. For this reason, I refer to this perspective as the 'intervention perspective'.

In contrast, designing theories of change from a 'system perspective' entails first considering a change of interest, and then figuring out what may have caused it: the intervention but also other contributing factors.

In my experience, the intervention perspective is simpler, more often used, but often rather useless. The system perspective, on the other hand, is more complicated, less often used, but often very useful.

In what follows, I describe both perspectives in more detail.

Intervention perspective

One way to approach theory of change is what I call 'forward explanation'. In this perspective, the intervention features as the root cause of subsequent change: starting from an intervention, the theory of change explains what immediate result(s) the intervention contributes to. Then, continuing along the causal chain, it explains what subsequent result(s) the immediate result(s) contribute to, and so forth (Figure 3.3). In an earlier publication (Belcher & Palenberg, 2018), Brian Belcher and I referred to this as 'intervention perspective'.
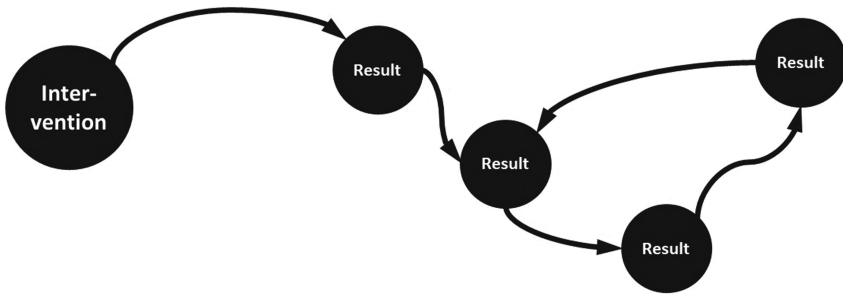
Figure 3.3 Example of a 'forward explanation' theory of change

The implicit, initial theory of change shown in Figure 3.1 is an example of an intervention perspective theory of change. It starts with the intervention of making results information available, which contributes to the intermediary result of staff using that information, which then helps the organization.

Theories of change designed from this perspective build on concepts like 'results chains' or 'impact pathways'. In this perspective, possessive language (such as, 'the results of the intervention') is often used. This type of intervention-centric language is widely used in evaluation literature. For example, the OECD Development Assistance Committee's Glossary of Key Terms in Evaluation and Results-Based Management defines results as the 'output, outcome or impact of a development intervention' (OECD, 2010), and the evaluation criterion of effectiveness is defined as the 'extent to which the intervention achieved, or is expected to achieve, its objectives, and its results' (OECD, 2021, italics added).

System perspective

Another way to approach theory of change reflects a 'system perspective' (Belcher & Palenberg, 2018). It uses 'backwards explaining' of an observed or desired change by means of its contributing causes, as shown in Figure 3.4. The key difference to the intervention perspective is that the intervention has—*a priori*—no special status among the contributing causes.

In this perspective, all relevant causes that contribute to a specific change should be described. Here *relevant* means that these causes—taken together— explain most of the change.

The theory of change in Figure 3.2 is a system perspective theory of change. It models how a small part of the ministry's evidence-based learning and decision-making system works and what factors contribute to its performance.
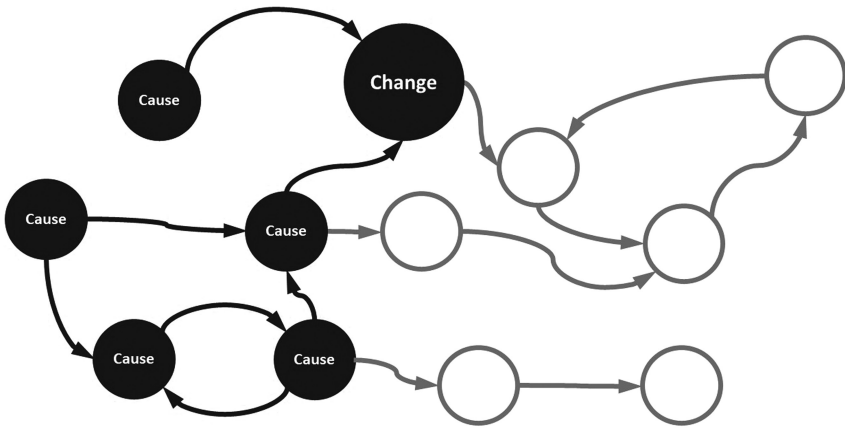
Figure 3.4 Example of a 'systems perspective' theory of change

Which perspective works better?

The intervention perspective is widely used. I believe this is because it reflects how we intuit simple cause-and-effect relationships in our daily lives. We say, for example, that we reduce our headache by taking an aspirin.

When we apply this type of intuitive monocausal thinking to more complex interventions and their effects, the result is an intervention perspective theory of change like the ones depicted in Figures 3.1 and 3.3.

Unfortunately, the validity range of intervention perspective theories of change is rather limited, making them ill-suited for answering the more complex questions of interest in the evaluations I have been involved in. Or, to keep it simple—but still beyond what an intervention perspective theory of change can handle—the question of how interesting you find reading this essay likely depends on many other causal factors beyond the fact that an aspirin reduced my headache while writing.

Strictly speaking, intervention-centric theories of change only serve their modelling purpose well when they describe cause-and-effect relationships that are indeed almost monocausal—that is, in which the intervention alone is the principal and dominant cause of the next result, which again is the principal and dominant cause of the next, and so forth.[2] However, in this special case, both perspectives will produce the same result.

---

2 This also applies if there are multiple root causes (for example, if different intervention activities are separately modelled) and if they contribute to more than one result at a time.

In all other cases, intervention perspective theories of change are prone to ignoring important contributing causes beyond the intervention. This is because they are guided by the question, How effective was my intervention? which risks putting blinkers on the evaluator's eyes.

For example, if we had conducted the entire evaluation referenced above of learning from results based on the initial understanding that the availability of results information represented the primary driver of all further change (Figure 3.1), then we could have easily missed out on the important elements of capability, opportunity, and motivation, which, in the end, turned out to be a good description of key enabling factors for the desired behaviour change. The risk of missing out on important causal factors would not go away if we had reflected them in assumptions on the side because the intervention-centric theory of change would still suggest that making results information available represents the principal driver of results-informed learning and decision making.

In contrast, system perspective theories of change are guided by the design question: What contributes to this change? They are therefore less likely to miss out on important contributing causes. We experienced this when we moved from the implicit, initial theory of change in Figure 3.1 to the more informed theory of change in Figure 3.2.

The intervention perspective theory of change in Figure 3.1 started with the silent assumption that the provision of results information would trigger its use. There might be other conditions and assumptions, but the causal relation between making the information available and using it would be the gist of it. When looking at what really contributes to results-informed learning and decision making, we quickly realized that this initial model was too limited and that we were missing out on other important factors. We therefore took a system perspective and started from the desired (behaviour) change: staff making use of results information. Then we checked the literature and found a simple model for explaining behaviour change. Suddenly, there were several additional candidates for contributing causes to that change. Importantly, the system perspective forced us to widen our focus from the single activity of making results information available to a wide range of other activities targeting any box and/or connection in the theory of change in Figure 3.2.

The relative importance of these other activities in contributing to results-informed learning and decision making was high. When we finalized our evaluation work, only one of seven conclusions and only two of eight recommendations concerned the availability of results information.

What does this tell us? For me, it primarily serves as a stark warning against oversimplifying theories of change.

In virtually all cases in my work, observed or desired change cannot adequately be described by a chain of monocausal cause-and-effect relationships. Instead, describing change as a consequence of multiple interacting causes—some of

which are related to the specific intervention at hand while others are not—often does the job.

Developing useful theories of change therefore requires using a system perspective. Despite much evaluation literature using intervention perspective language—such as 'the intervention's results'—we must consciously avoid ignoring other contributing causes because they often turn out to be more relevant and important than the ones we initially had in mind.

References

Belcher, B., & Palenberg, M. (2018). Outcomes and impacts of development interventions: Toward conceptual clarity. *American Journal of Evaluation*, *39*(4), 478–495. https://journals.sagepub.com/doi/10.1177/1098214018765698

Mayne, J. (2019). *The COM-B theory of change model*. Working paper. An earlier version of this paper can be retrieved from www.researchgate.net/publication/323868561_The_COMB_ToC_Model4

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*, 42. www.implementationscience.com/content/6/1/42

OECD. (2010). *Glossary of key terms in evaluation and results-based management*. OECD Publications.

OECD. (2021). *Applying evaluation criteria thoughtfully*. OECD Publishing. https://doi.org/10.1787/543e84ed-en

Palenberg, M., Bartholomew, A., Mayne, J., Mäkelä, M., & Esche, L. (2019). *How do we learn, manage and make decisions in Finland's development policy and cooperation: Management of results information and knowledge at the Ministry for Foreign Affairs of Finland*. Final Report. Ministry for Foreign Affairs of Finland. Retrieved from https://um.fi/publications/-/asset_publisher/TVOLgBmLyZvu/content/evaluointi-tietojohtamisesta-miten-opimme-johdamme-ja-teemme-paatoksia-suomen-kehityspolitiikassa-ja-yhteistyossa-/384998

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press.

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic Books, Hachette Book Group.

4 Theories of change between critical thinking and social practices

Peter Dahler-Larsen

Programme theory and critical thinking

Using a programme theory is one of the best ways to give critical thinking a prominent role in evaluation. Theory requires the explication of critical assumptions about how interventions can be transformed, step by step, into outcomes that are socially and politically desirable. When the logical status of each of these steps is made clear—some are necessary, some are sufficient, some are very specific for the intervention at hand—then logically strong inferences can be made about the causal link between an intervention and its outcomes, even if the available empirical material is limited.

However, our daily lives and our organisational practices are not ruled by logic alone, to put it mildly. Our cognitive capacity is limited (Simon, 1996), we use shorthand cognitive heuristics (Kahneman, 2011), we use rough and inconsistent social typifications (Schutz, 1973), we make normative assessment intuitively (Sayer, 2011), and many of our interactions in and between organisations are based on normative expectations and institutional scripts (Dimaggio & Powell, 1983). Thus, our everyday evaluative practices take place in the grey area between strict logical thinking on the one hand and social practices on the other. As suggested by Ling (Essay 5), programme theories are tools we use to navigate an unknown future. They are not meant to be accurate pictures of reality.

In this essay, I discuss three related aspects of the use of programme theories in this grey area: (1) the asymmetrical relation between confirmation and falsification of a theory, (2) the use of theory for (primarily) presentational purposes, and (3) the balance between monologue and dialogue in programme theory. I shall use *programme theory* as a generic term for the causal logic that explains how an intervention is supposed to work, while I use the term *theory of change* when that term is used in the specific example or argument at hand. Although there may be differences between the two terms in some contexts, the argument in this article applies to both.

1. The asymmetrical relation between confirmation and falsification of theories

I recently proposed an evaluation of a whistleblower policy. I suggested an empirical focus on what the literature on whistleblowing identifies as major threats to effectiveness: lack of knowledge about the policy, lack of trust in independent and objective complaint management, lack of protection of anonymity, and fear of retaliation. One of the potential funders (a national umbrella organisation) said that it did not want a negative focus. They wanted the evaluation to prove how whistleblower arrangements could operate optimally (to the benefit of their member organisations). Many factors explain why we did not come to an agreement. One of them is the fundamentally asymmetrical relation between falsification and confirmation of theories. Critical argumentation and institutionally sustained hopes do not go together well.

Since Popper (1959), we have known that it is almost impossible to confirm a theory. What we can do, however, is subject it to a number of tests, and if the theory does not come out as falsified, we continue to have trust in it. Given this logic, we can actually falsify a simple little theory based on a limited empirical material. If we have a theory that says that X leads to Y everywhere, and we have just one case of X without a Y, the theory is falsified. So, contrary to what some believe, it is possible to produce some causal statements even with a limited empirical material and even in the absence of a counterfactual.

This philosophical point is more relevant to evaluation than many believe. If we ask the question 'Does X causally lead to Y', we can answer the question by posing a programme theory predicting how that would unfold, with all the steps explicated, and then check whether all these steps occur or not. But what convinces us most strongly is if we find that a necessary step between X and Y is absent. Then we know that either there is no link between X and Y or the link in the programme theory is not correctly specified. No matter what, the programme theory is falsified in the case at hand, which is equivalent to saying that X does not work as expected (in the case at hand). Needless to say, an absence of Y itself also nullifies all causal statements saying that X produced Y in the case at hand. It is more burdensome or more rare to find compelling evidence that proves that Y is there, and that it was unequivocally caused by X.

While the philosophical arguments are favourable to a falsification of theories, the dominant practical and social forces often pull in the other direction. Foundations finance projects based on trust in official programme theories. Even if foundations claim they are open to learning through trial and error, grantees may be unwilling to take steps in that direction. A grantee may feel it unwise to share too many thoughts about threats to validity of the programme theory. In addition, since grants are given on an ad hoc basis, the grantee may also need other grants from other foundations in the near future. Therefore, it might be safer to continue to use an official programme theory for presentational and

promotional purposes rather than as an instrument in systematic and critical causal analysis.

2. The use of theory for (primarily) presentational purposes

By the use of a theory for presentational purposes I mean that it is mostly used as a communicative device to signify that there is an orderly relation between inputs, processes, outputs, and outcomes, at least graphically. Let me give an example. Since 2018, a Danish philanthropic foundation supporting innovative social programmes has asked all its grantees to present the core idea of their project in the form of a theory of change. The foundation also sets aside about 10% of the donation for external evaluation of each project. In a recent review of their evaluation practices, I read all the evaluation reports subject to these requirements. They all provided clear, readable, and understandable graphic representations of their theory of change in an early chapter in the evaluation report. I found that the theory introduced outcome measures that were actually used in the subsequent data collection. However, I also found that the full potential of theory as an integrated tool in all aspects of the evaluation process was far from reached. The reports did not explain how the theory of change was delineated, so the reader knows little about whether it speaks to, say, existing theory in the field or to particular stakeholder expectations, and if yes, which stakeholders.

The theories of change were not logically linked to a particular understanding of the social problems that the interventions were meant to address.

While data were collected that corresponded to the outcomes stipulated in the programme theories, little attention was paid to all the processes, step by step, which could have been used to qualify a causal statement about whether the intervention was responsible for the changes in outcomes. Therefore, the capacity of the theory as a tool for process-oriented causal analyses was not in operation.

After the empirical data were collected and analysed, it was rare to find consequences drawn for the theory, so the reader was left without a clear sense of whether the theory should be confirmed, revised, or rejected as a result of the evaluation.

In no instance did I find a new, updated version of the theory of change, which could have been a neat, logical stepping stone for the next generation of projects in the same policy area. When recommendations were made, they were not clearly rooted in a theory, old or new.

In sum, in my review I found that the theory of change played a role in giving an overview of the project, and it provided an orderly presentation of the logical link between interventions and outcomes. It also helped identify what counted as relevant outcomes. However, it was not used throughout the evaluative process as a tool for critical thinking. It was not visible how it might have contributed to qualified causal inference. For example, the reports did not engage in logical

arguments about which steps were seen as necessary or sufficient for outcomes, nor attributable to the intervention at hand.

Many factors may contribute to this situation. Grantees may feel that their grant is based on the presented theory of change, so the value of their project is undermined if the theory is challenged. They may also fear that the foundation does not want to hear bad news (even if the foundation is, in fact, clear about how it accepts trial and error as a way to learning). The external evaluators face incentives that support a short positive report rather than a long rigorous and critical one. In addition, the organisations that receive grants may be small and partly based on voluntary work, and all their focus is on the substance of their social work—not on evaluation capacity building. On that basis, perhaps they believe that a continued belief in the theory of change already presented is best for their reputation and for their chances to receive additional funding, maybe from other sources. If they believe so, experience actually supports them. They have already received funding exactly because of the very same theory. Perhaps they think: Don't fix it if it ain't broken. If it secures funding, it is not broken, so do not rock the boat. In addition, grantees may also think that the theory of change itself is a neat representation of their project. It may be difficult to depart from it if it represents the very identity of the project.

If grantees think like that, it would be quite logical—with logical here referring to their organisational history and institutional position as grantees and providers of human services that are notoriously difficult to turn into successes (Hasenfeld, 1983). To maintain faith, they have to believe in the logic of what they do.

3. Programme theory: From monologue to dialogue

In my review of evaluative practices among grantees, I found that once a programme theory was represented, it was not changed. No evaluation report provided a modified or refined version as a result of the evaluation. There is no sign of 'talking back' to the programme theory—not from participating stakeholders, not from data, and not from the evaluator.

This suggests that there is a potential in using programme theories in a more dialogical way. Such dialogue may not be conventionally conceived as a contribution to rigorous testing of causal statements. Nevertheless, dialogue can contribute to an exploration of interesting and important causal pathways that may lead to policy change. Let me give a final example.

In an advisory group on evaluation of national tests, we discussed the underlying test principles. In linear tests, all students get the same test items in the same order. In adaptive testing, a computer algorithm determines the choice of the subsequent test item based on whether the student answered the previous test items correctly or not. Experts believe that since an adaptive test iteratively adapts test items to each student, it would determine the proficiency level of each student more accurately.

However, in the most recent evaluation, teachers actually complained about the existing adaptive tests. Teachers found it impossible or burdensome to understand what the tests were about and how to use the results in daily teaching. Some members of the advisory group argued that the existing adaptive test principle was responsible for this problem. Since every child receives their own combination of test items, it becomes difficult for a teacher to get an overview of the performance of a whole class or groups of students. Adaptive tests also lead to test items that are perceived as extremely difficult and sometimes meaningless (since the underlying computer programme has to guess the proficiency level of each student by iteratively giving too easy or too difficult items).

So, through the deliberations, the advisory group concluded that although the adaptive principle might be technically superior, it was also causally responsible for the widespread perceived uselessness and meaninglessness of the existing tests. All things considered, the advisory group recommended linear instead of adaptive testing. The politicians adopted the recommendation in a new national test system.

The example shows that under some conditions, an open deliberation about potential causal pathways can help facilitate policy change (even if, in this example, there was no strict causal analysis, merely a compelling argument). The combination of the views and experiences of different stakeholders into the same causal model sometimes leads to fruitful insights, even if the resulting model is still fairly simple. For example, it can be found that one variable has several causal consequences, so a phenomenon that is justified in an official programme theory may impede positive change according to alternative programme theories (Dahler-Larsen, 2018). This is what happened in the deliberations about adaptive testing.

Perhaps the deliberation was successful in this case only because the national tests were already subject to intense public debate and because the collaborative process in the advisory board provided a particularly fertile ground for deliberation about multiple causal pathways, not just a test of the conventional beliefs.

A fuller exploitation of the potential in programme theory hinges on the institutional, political, and organisational conditions in which it is embedded. Theory-based evaluation is promising. More can be done with it. *How* much more depends less on the logic of causal analysis itself than on how the evaluative process is socially constructed in interaction with organisational and political factors. There has to be room for negative findings and alternative interpretations. And enough trust in the fruitful consequences of going into that room together.

References

Dahler-Larsen, P. (2018). Theory-based evaluation meets ambiguity: The role of Janus variables. *American Journal of Evaluation*, *39*(1), 6–23.

Dimaggio, P., & Powell, W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(6), 147–160.

Hasenfeld, Y. (1983). *Human service organizations*. Prentice-Hall.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Popper, K. (1959). *The logic of scientific discovery*. Hutchinson.

Sayer, A. (2011). *Why things matter to people: Social science, values and ethical life*. Cambridge University Press.

Schutz, A. (1973). *The problem of social reality: Collected papers*, Vol. 1. Nijhoff.

Simon, H. A. (1996). *The sciences of the artificial*. Massachusetts Institute of Technology.

5　How deep is your ontology?

How ontological thinking can improve how evaluators use theories of change

Tom Ling

Introduction

Theory-based evaluation is a creative and dynamic part of current discussions of the intersections among evaluation, implementation, and complexity (Mayne, 2017; Mickwitz et al., 2021; Stame, 2022). It should be emphasised that all evaluation is theory based to one degree or another. The UK government's guidance on evaluation defines *theory-based evaluation* as: 'can be used to investigate net impacts by exploring the causal chains thought to bring about change by an intervention' (HM Treasury, 2020, p. 43). It is an approach that can in principle be combined with any methodology (theory-based evaluation is methodologically agnostic), but in practice where comparison is possible between an intervention group and a comparator group the preference is most often not for theory-based evaluation but for experimental or quasi-experimental methods. At heart, theory-based evaluation involves developing a narrative that can be supported or challenged by evidence about how the intervention is intended to work and how it works in practice.

A theory of change is key to building this narrative. It is an artefact that has considerable value in the process of evaluating complex interventions:

- It enables stakeholders to contribute their own understanding and better engage with others' views.
- It provides a good first take on data collection.
- It surfaces assumptions (and some assumptions do not survive such a reality check).

However, alongside these benefits, it brings risks. Reflecting on the nature of the social world—where interventions are delivered, and evaluations are constructed—can greatly help to understand, and then manage, these risks. It can also help identify a special set of circumstances where a theory of change should not be used, or at least used very differently.

A theory of change has been described as a 'description of a sequence of events expected to lead to a particular desired outcome' (Davies, 2012). Like a map, it represents things of importance—on a map this might include roads,

bridges, cliff edges—and how they connect to each other. Like a map, a theory of change has value because:

- It abstracts from the detail of everything and helps make sense of the key events through which it is hoped to achieve results.
- It can guide the evaluator towards which data to collect and away from data that are merely 'nice to have'.
- It can highlight for decision makers potential problems that had not been properly considered or opportunities to exploit.
- It can clarify where resources are to be put and behaviours changed and allow a conversation of who needs to do what.

All this can be shared across evaluators, implementers, and stakeholders whose worlds are touched by the programme. Theory-based evaluations use the theory of change as a foundational step in developing their evaluation framework. As noted by Palenberg (Essay 3), the theory of change is a qualitative conceptual tool that helps us to systematically analyse and understand potential cause-and-effect relationships.

Using a rich ontology to manage the risks of theories of change

Gasper identifies potential problems with theories of change:

> 'logic-less frames', where only an illusion of logic is provided; 'jamming' of too much into one diagram; 'lack-frames', which omit vital aspects of a project; and 'lock-frames', whereby programme learning and adaptation are blocked. (Gasper, 2000, pp. 21–22)

As evaluators, we are often aware of these problems but perhaps rarely reflect on what this means for our practice. In response to these problems, I propose five principles, or propositions, that should underpin the work of evaluators when using theories of change. This discussion is relevant for all evaluations using a theory of change, but it is especially relevant for evaluations of more complex interventions.

The characteristics of the social world that should be considered when using a theory of change derive from the work of Lawson (2003) and are paraphrased here:

1. The future is always to some extent *open*. Human agency and judgement (for example, in learning, sense making, and adapting) are not predetermined. However, the intervention is not the only thing shaping events and…
2. …the future is also *structured* by power, gender, economics, and organisational imperatives that shape how programmes work (and possibly not in ways anticipated in the theory of change) and (especially) how people exposed to

interventions react. However, although structured, in causally dense interventions the consequences are not fixed, and we can anticipate that change measured at any one moment in time using any one academic lens will not be permanent because…

3. …social, psychological, biological, and other processes combine to create *emergence* and *dynamic* outcomes that not only change over time but may also be hard to predict because…

4. …systems in the social realm not only have tipping points involving sudden change, but also social systems are *internally related* to each other with multiple overlapping systems combining to create change (for example, gender, race, economics, and politics). Any intervention in one system may create related responses in others, with potentially cascading and unfolding changes but, perhaps more often, also resistance that undermines intended changes leading to a reversion to the norm. Fully independent systems in the social sphere are rare or non-existent.

It should be apparent that a theory of change is not an ontological statement. It is not designed to define reality or to capture the complexity of the world but, rather, to provide a means to manage that complexity. Evidently the real world does not conform to a logic model (which is a deliberate abstraction from the real world for the purpose of better understanding that world). However, a theory of change can be an invaluable first step in an evaluation (or in developing and improving a programme[1]).

A theory of change identifies the sites where individuals and organisations will be exposed to the programme (both as part of delivery processes and as users or consumers). This allows evaluators to design an evaluation that will understand how choices, habits, and systems interact with the programme design under conditions where individuals are uncertain about causality but nevertheless must act. To put it less grandly, humans muddle through with incomplete information and incomplete causal understanding. Important evaluation questions include understanding critical leverage points driving how (well) programmes work: how people and organisations became aware of the programme, how they understand it, their available resources to engage and act, and their ability to mobilise others in response to the programme. The role of the evaluator is to explore these sites and assist decision makers in identifying if, when, and why they produce the consequences anticipated in the theory of change. In this sense, every theory of change-based evaluation is not simply an exercise in 'proving' or 'disproving' what is known about the theory of change; it also supports learning by doing. An ontologically informed use of theory of change provides a pragmatic tool for structuring social learning, explaining why results vary over time and by context, and thereby strengthens our ability to understand and act in a complex world. Using theories of change, we most often chart programme

succession and evolution rather than programme success; no programme stays the same—far less succeeds—forever.

Evaluating transformative programmes

Programmes differ in important ways. Some programmes simply aim to do more, faster, cheaper, or better. But fundamentally they ask people to do things that sit within their existing experiences. From previous evidence and prior experience, implementers can reasonably be expected to make assumptions about how people and organisations might react. Other programmes, however, invite individuals and organisations involved with the programme to experience something new. Famously, Henry Ford is supposed to have said: 'If I'd asked my customers what they wanted, they would have said "a faster horse"'. Innovative programmes invite you to drive a car and not simply ride a faster horse. At the start, neither participants nor evaluators can be sure what this will be like or how it might best be measured. This is where the theory of change as a tool for exploring the unknown becomes valuable.

Especially in innovative or transformative programmes, if we confuse a theory of change with the real world, the evaluation becomes seriously unstuck. Our four characteristics of the social world outlined previously identify areas of uncertainty about how any programme might work. The farther we move from the 'faster horse' model of change, the more the future becomes uncertain; structures such as gender, racism, and inequality will be relevant but may play out in ways that are hard to anticipate and events might accelerate towards unexpected (and unwanted) outcomes. The less the future looks like a faster horse, the more the evaluation needs to be part of a process of discovery, working alongside those implementing the programme and those who are exposed to it to understand how the programme works, for whom, and in what circumstances. Delahais (Essay 1) makes this point in relation to achieving sustainable forest management, for instance, and provides important insights to how to structure this context in a theory-based evaluation. Doing so requires us to understand both the assumptions implicit in the theory of change and whether these assumptions prove to be well founded.

In this context, I suggest five pragmatic principles to apply when using theories of change (and especially for innovative programmes):

5. **Engage stakeholders.** Understand how individuals and organisations engage with the programme and what they seek to achieve by participating. This may not be how the programme implementers anticipate it. In complex and innovative programmes, neither the programme designers nor the evaluators can accurately predict the behaviours of those in the programme. Stakeholders also provide unique insights that help understand systems, structures, and emergence from the perspective of those living within the settings where

people are exposed to the programme. They are well placed to assess the plausibility of the theory of change, potentially dismissing implausible causal pathways implicit in the design of the programme (Ton, 2021). For this reason, qualitative data are essential partners to quantitative data to understand how social actors reflect upon, learn in, and make sense of the circumstances in which they react to the programme. Stakeholders are well placed to explain how these systems might work and to understand how the 'real' logic of the programme might differ from the description in the theory of change (although, of course, stakeholders are not infallible). Palenberg (Essay 3) recommends using COM-B model for behaviour change to structure understanding of the drivers of behaviour among those who engage with a programme—addressing how behaviour might be explained by three groups of conditions: capability, opportunity, and motivation (Michie et al., 2011).

6. **Recognise structural and systemic constraints.** These may influence outcomes as much as the activities identified in the theory of change; evidence for this will be found in the patterning of outcomes (especially when these patterns were not intended and involve the system returning to 'business as usual' after the effects of the intervention dissipate). For example, programmes designed to reduce structural inequalities may unwittingly reinforce or displace these inequalities regardless of the content of the theory of change. Equally, behaviours may often adapt quickly to a programme only to later settle back into their previous systemic ways of working. Sustainable change will ensure the structures and systems are transformed to support the theory of change.

7. **Iterate the theory of change.** Be sensitive to understanding that changes in the outcomes of interest may progress or reverse in waves or tipping points. As systems become more apparent, and as results ebb and flow, each version of the theory of change should be provisional. Theories of change are artefacts to support social learning, and they should help stakeholders who are muddling through and still attached to prior assumptions, while finding structural and systemic constraints confusing since change is lumpy and omnidirectional rather than linear.

8. **Be theory led.** Focus on understanding causal pathways as they evolve in their social contexts. Using theories of change, and in general drawing on theoretical models, provides evaluators with opportunities for mid-range theorising and reflection and step-back opportunities. However, Dahler-Larsen (Essay 4) reminds us that there is an asymmetry involved in testing these mid-range theories; we may be better able to disprove a theory than prove it.

9. **Be humble and curious.** Evaluators are part of the uncertainty of a complex world and, while the theory of change is a helpful tool for managing this uncertainty, it is a tool best used *with* stakeholders and not *on* them. Evaluations are events in complex systems, and evaluators—like all social actors—are attracted to prior assumptions, constrained by experience, and

work with an incomplete understanding of the world. Like all other stake-holders noted in Dahler-Larsen's essay, evaluators' cognitive capacity is also limited (Simon, 1996), and we rely upon cognitive heuristics (Kahneman, 2011).

Why a little ontological reflection is helpful

Ontology is concerned with the nature of reality, of being and becoming. It is often considered alongside epistemology, which is concerned with how we know that reality. The following paragraphs explore: (1) how theories of change always include implicit assumptions about ontology, (2) how these assumptions are often not well founded (and evaluations are typically not resourced to explore them), (3) how some ontological reflections can help address this (even within the resources often allocated for evaluation), and (4) how best to respond to this. Among other things, this essay will argue that evaluations should always see social actors involved in a programme as people who have agency and creativity, who learn and have insights about the systems they live within, and who are not simply the carriers of a programme logic. Working iteratively and drawing upon the situated learning of different stakeholders, an in-depth understanding of 'social facts' can be developed. This helps understand how well the assumptions contained within the theory of change are matched by the realities of the social world. The purpose certainly is not to advance the philosophy of ontology but rather to improve the practice of working with theories of change by reflecting ontologically.

Theories of change need to be simple enough to be a useful 'map'[2] by identifying only the most important elements in the pathways, while avoiding being misleadingly linear and sequential. Furthermore, a useful theory of change will not only map the activities and observable events along a sequence of events but will also suggest where to look for the less observable and less tangible mechanisms that connect these events—such as trust, shared values, and identities. Levine (Essay 2) makes this point in relation to the challenging task of evaluating impact in relation to policy change. These events provide a focus for the evaluation. However, to make sense of them, their context needs to be understood, ranging from (for example) the design and location of buildings and rooms where the events take place, to the public financial systems resourcing the programme, to the power of professionals delivering the service, or to the rise of consumerism among service users. Ignoring these contextual factors may lead to an exaggerated sense of what outcomes can be attributed to the programme or, conversely, a misunderstanding of where the causes of any failure lie.

Therefore, theories of change are consciously and deliberately ontologically 'thin'—they provide a one-page, high-level description of highly complex activities. The solution is not to 'fatten' them up but to use them with care. A

theory of change is an artefact to aid speculation about future events or a basis for reviewing such speculation after the event. It is not a statement about the real world but a means to explore how the world might be (both by acting differently in the world and by evaluating these actions).

Although theories of change may be ontologically thin, they can be helpful stepping stones to understanding the world within which the programme is an event. The theories of change contain theories about this world that can be tested. These theories are often naïve, including simplistic models of human behaviour from economics that see people as simply responding to incentives to maximise their personal benefit (utility). These models are misleading and the epitome of a thin ontology. As Stame (2004) asserts in relation to realist evaluations: 'So the mystery of the black box is unveiled: people inhabit it. This makes for a completely different design of evaluation' (p. 62). Theories of change do not 'do' anything; we use theories of change to structure our understanding of what people do, and why, and with what consequences.

Building on theories of change for added depth and understanding transformation

One risk of leaving people out of the black box is that when results are reported, evaluations take an overly narrow view of the evidence that explains performance. A programme is an event in a world that is already structured and dynamic and nothing that happens is an outcome of the programme alone. One often-noted consequence of this error is that when others try to replicate an approach that has apparently worked, they achieve very different results or they quickly find themselves having to adapt the approach but with insufficient information from the evaluation.

Increasingly there is a drive towards transformational change or sociotechnical transitions in response to potentially catastrophic challenges of climate change, social fragmentation, ecological degradation, and so forth. Evaluators, to continue to be useful, need to adapt their existing tools and develop new ones, especially in relation to transformational change where social learning and adaptation are key to success.

According to Patton (2020): 'evaluating transformation is different from evaluating projects and programmes' (p. 22). A particular kind of theory of change is required of a programme intending to achieve a transformation in car use, for example, when transitioning to electric cars (Sorrell, 2015). The global car industry, car-servicing suppliers, regulations, the built environment, the petrochemical industry, lobbying groups, and car culture all interact to create historically rooted structures that make it profoundly difficult to deliver transformation (Sorrell, 2018). Changing them requires not only an ability to both understand and work with systems and how they have emerged but also to mobilise and inform human creativity and agency.

Theories of change are necessarily superficial descriptions of a desired sequence of events that can helpfully direct attention or help structure conversations. I emphasise the value of enriching this necessary superficiality by emphasising the importance of human agency, learning, and sense making. However, I also concur with Palenberg whose essay emphasises the importance of bringing in a system perspective. As individuals and organisations engage with the transformational programme, how will they adapt their views and change their behaviours as new information comes to light? How will existing organisations adapt to new pressures and opportunities as systems change and relationships evolve? Will entrenched power relationships evolve smoothly towards more inclusive and socially just results? The theory of change can be visualised as just the tip of the iceberg, visible above the water, while the underlying assumptions make up the much larger part of the whole but are less visible below the water line. The drivers of change lie beneath the surface.

Finally, it is quite possible that there are circumstances under which we should not try to use a theory of change in a conventional way. Its use should not be to evaluate a planned change but, rather, to enrich a process of discovery. Returning to our four ontological insights, these circumstances are when: (1) the future is highly open and there is no body of research or experiential knowledge to predict outcomes with any certainty, (2) the social and economic structures are yet to establish themselves, (3) the world is dynamic with as yet limited understanding of patterns across time, (4) and systems emerge and interact in ways that are so unpredictable that linearity is as possible as nonlinearity. These circumstances do not characterise most evaluations but in relation to some of the major challenges of the age, they are more common.

Conclusions

To summarise: a rigid application of (even a well-designed) theory of change to evaluate complex interventions has at least two risks highlighted here: (1) being 'futures illiterate' (that is, blind to key uncertainties) and (2) ignoring the role of individual and organisational social learning and agency. Without considering these risks theories of change:

1. Exaggerate our ability to anticipate future conditions and reward 'sticking to a failing plan' over learning and adaptation. It therefore limits the interventions' outward-looking cognitive engagement with a changing world that supports adaptation.
2. Underplay the importance of choice-making 'on the fly' and therefore (if enforced) would undermine the sense of individual and organisational agency and creativity upon which sustainable transformational change depends.
3. Create misleading assumptions about how easy it might be to spread and scale up interventions deemed to be successful by evaluators.

There are also many reasons why, after taking stock of what is possible and what has been learned, people's preferences change and part of the programme logic is abandoned in favour of a route that leads to different (and, some may think, better) outcomes. In certain circumstances, the evaluation should focus on arriving at a judgement of the worth of discovery. What is important is how well founded these choices were, and what principles and values were invoked in making them. Were changes in direction the result of careful and balanced judgement or of something else (such as a new 'flavour of the month', change of personnel, or organisational crisis)?

In conclusion, I am arguing for using theories of change more flexibly and deeply rather than abandoning them. Understanding and sharing the programme logic remains important. A theory of change is never a solution to understanding the programme logic, but it may be (and often is) a very helpful way of framing how we communicate about these causal pathways. It also helps engage with stakeholders in discussing how the causal pathways relate to a successful evaluation (and, indeed, intervention). In complex evaluations it helps structure conversations among stakeholders over time and to track evolving expectations and approaches. I suggest adopting the five pragmatic principles for an ontologically informed approach.

Notes

1 The term *programme* is used to describe the interventions under consideration in this essay, but this is meant to include a range of interventions and approaches.
2 In 1946, Jorge Luis Borges famously described an empire where 'the cartographers guilds struck a map of the empire whose size was that of the empire, and which coincided point for point with it. The following generations, who were not so fond of the study of cartography as their forebears had been, saw that that vast map was useless…'. This is reproduced in Borges (1998, p. 325).

References

Borges, J. L. (1998). On exactitude in science. In *Collected Fictions*. Translated by Andrew Hurley. Viking Penguin.

Davies, R. (2012). Criteria for assessing the evaluability of a theory of change. *Rick on the Road.* Retrieved from www.mande.co.uk/wp-content/uploads/2012/Evaluablity%20of%20TOC%20criteria.pdf

Gasper, D. (2000). Evaluating the 'logical framework approach'—towards learning-oriented development evaluation. *Public Administration and Development*, *20*(1), 1728.

HM Treasury. (2020). *The magenta book: Central Government guidance on evaluation.* HM Treasury.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Lawson, T. (2003). Ontology and feminist theorizing. *Feminist Economics*, *9*(1), 119–150.

Mayne, J. (2017). Theory of change analysis: Building robust theories of change. *Canadian Journal of Program Evaluation*, *1*(2), 155–173.

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*, 42. https://doi.org/10.1186/1748-5908-6-42

Mickwitz, P., Neij, L., Johansson, M., Benner, M., & Sandin, S. (2021). A theory-based approach to evaluations intended to inform transitions toward sustainability. *Evaluation*, *27*, 281–306.

Patton, M. Q. (2020). *Blue marble evaluation: Premises and principles*. Guilford Press.

Simon, H. A. (1996). *The sciences of the artificial*. Massachusetts Institute of Technology.

Sorrell, S. (2015). Reducing energy demand: A review of issues, challenges, and approaches. *Renewable and Sustainable Energy Review*, *47*, 74–82.

Sorrell, S. (2018). Explaining sociotechnical transitions: A critical realist perspective. *Research Policy*, *47*(7), 1267–1282. https://doi.org/10.1016/j.respol.2018.04.008

Stame, N. (2004). Theory based evaluation and types of complexity. *Evaluation*, *10*(1), 58–76.

Stame, N. (2022). Program, complexity, and system when evaluating sustainable development. *Evaluation*, *28*(1), 58–71. https://doi:10.1177/13563890211065488

Ton, G. (2021). Development policy and impact evaluation: Learning and accountability in private sector development. In H. Zafarullah & A. S. Huque (Eds.), *Handbook of development policy*. Edward Elgar Publishing.

Part 3

Using theories of change to understand wider societal change processes

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

6 Theory of change for sustainable business

Jens Andersson

As social and environmental challenges mount globally, from the accelerating effects of climate change and increasing loss of biodiversity to the growing economic and social inequalities, there is increasing pressure on the private sector to contribute proactively with solutions (Meyer, 2018). This goes beyond traditional corporate social responsibility (CSR) approaches that have been criticized as being vague, focused on philanthropy, and disconnected from business activities (Porter & Kramer, 2006; Rønholt Albertsen, 2021). Instead, the new corporate sustainability agenda involves transformation of core business models, which in turn requires companies to think more strategically about the interplay between their business and society.

This essay focuses on the role that the theory of change approach can play in this new and deeper corporate sustainability agenda. While corporations have been exposed to theories of change for specific interventions and partnerships within the CSR context for some time, little seems to be written about it. This indicates that the theory of change approach has not yet entered the core corporate strategic toolbox or been picked up by the business literature as part of the new and deeper corporate sustainability agenda. This essay contributes to filling that gap, based on my experiences with applying a theory of change approach within the global furniture company IKEA.

I argue that the theory of change approach is indeed an important tool for companies to understand and strategize within the business-society nexus, but for it to be used in this way, it needs to be adapted to a private sector context primarily driven by business logic. This requires pragmatic use of the tool itself and a learning mindset, while accepting that it may come at the cost of loss of rigour in the way theories of change have conventionally been conceived and used in a development context.

The emergence of sustainable business

There is no doubt that sustainability has increasing urgency for the private sector. International surveys show that companies' sustainability work has developed and deepened over time (KPMG, 2020; Scott, 2021). In step with increasing social and environmental challenges, corporate sustainability is driven by

consumers, legislators, and companies themselves, who see both increased business opportunities and risk mitigation potential. Meanwhile, the sustainability agenda is expanding to integrate complex areas, such as human rights, climate, circularity, and biodiversity.

These developments are forcing many companies to question established strategies, manage growing risks in their supply chains, experiment with new sustainable business models, and reconsider the connection between profit maximization and societal benefits (Bocken et al., 2014; World Economic Forum, 2019). Both path and goal are far from clear, and the need for innovation and new thinking in the private sector and society at large is enormous.

The rapprochement between business and societal goals and challenges is contributing to blurring the traditional division between commercially and socially oriented activities, instead creating a continuum between the two (European Venture Philanthropy Association, 2020). While traditional companies and investors are becoming more sustainable, at the other side of the spectrum traditional NGOs are adopting business thinking, and social enterprises (combining entrepreneurial activity with a social purpose) are proliferating within what the European Commission calls the 'social economy' (see, for example, European Commission, 2021).

A strategic tool for sustainability

Within the nascent, hopefully more sustainable, but complex, corporate world, the need is increasing to systematically understand companies' impact on society, integrate sustainability into their strategy work, and monitor and report on sustainability. In support of this shift, extensive international legislative and standardization work is underway from organizations such as the European Commission, UN Global Compact, and the Global Reporting Initiative, while other initiatives are working to more clearly integrate sustainability into accounting standards. Over time, these multiple, uncoordinated efforts have created a diversity of guidance, standards, and frameworks that observers have described as an 'alphabet soup' (Murray, 2021).[1]

In this plethora of emerging corporate sustainability frameworks, the theory of change approach is often proposed as the central tool to understand businesses' social and environmental impact. This is not surprising, given that it is particularly adapted to the formulation of goals and means in complex societal areas

---

1 See https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en (European Commission), www.unglobalcompact.org (UN Global Compact), www.globalreporting.org (Global Reporting Initiative), and www.ifrs.org/groups/international-sustainability-standards-board/ (International Sustainability Standards Board).

that are outside organizations' normal operations and control. Indeed, theories of change are already used by front-runner businesses with explicit social focus, such as impact investors that need to identify and select investment objects and social enterprises that need to formulate and monitor the social and environmental effects of their activities. In traditional commercially oriented companies, however, the theory of change approach seems to be used more ad hoc, if at all, despite it being part and parcel of corporate sustainability guidelines for some time (see, for example, WBCSD, 2013).

What it can and cannot do

In my work within IKEA, I have participated in applying a theory of change approach to a variety of contexts to formulate prioritized organizational, economic, social, and environmental goals and objectives: at the corporate strategy level; within corporate alliances, organizations, and initiatives with social and environmental objectives; for social enterprises; and to shape sustainability-oriented philanthropic programmes and partnerships.

This experience confirms the value of applying a theory of change approach—in both its structure and creation process—to support at least the 'pure' corporate sustainability agenda (for example, how a company should approach biodiversity). A theory of change approach allows for an inclusive discussion using the established output-outcome-impact framework. Taking a step back and holistically reflecting with peers or colleagues in this way seems to give relief from the action-oriented, jargon-filled, and vertical corporate context. At the same time, people new to the tool may have difficulties distinguishing between the different results levels and normally need external facilitation and hand-holding during workshops and later iterations to produce a clear and consistent theory of change.

Applying a theory of change approach to a mixed context, where there are both social and commercial goals, also adds complexity (see also Essay 7 by Penny Hawkins and Zazie Tolmer in this volume). Fundamental questions of how social and business goals relate to each other inevitably come up: Is this business with impact or business for impact? How business oriented can a social enterprise be without experiencing mission drift away from its social purpose? And vice versa: How much is a purely commercial company willing to adapt its business model and invest to become more impactful? And in business collaborations and partnerships between large companies and small social enterprises: How much do differences in size and power influence the theory of change design?

The answers to these questions are at the heart of the deeper corporate sustainability agenda and search for new business models. While sustainability is often touted as a business opportunity or risk mitigation strategy, it inevitably adds complexity, tension, and costs at least in the short run. Existing business

models are questioned and the potential benefits (for example, brand value) may not occur at the same time or place in the value chain as the costs (for example, increased purchasing costs). The ultimate outcome in terms of the commercial and sustainability mix depends on many factors, such as the commitment of top management and owners, external pressure, the type of industry or business model, internal capabilities, incentives, customer response, systemic pre-conditions, everyday decisions by middle-management and staff, and the list goes on.

This complexity means that the formulation of a theory of change, however refined, will not provide immediate answers, but it is a good tool for starting a conversation around the context and factors that need to be considered and formulating a common way forward. Applying a theory of change approach helps to spell out and deal with complexity and tensions. But it inevitably will be a journey, as there will always be goal conflicts and recognition that commercial enterprises—regardless of their sustainability ambition—ultimately need to be profitable to prevail and prosper.

Promoting its use

If it has such potential, how can the theory of change approach be promoted within a corporate context? This is by no means automatic, particularly in commercially focused companies. There may be resistance to the tool itself; it may appear foreign and complex. The structure and creation process may be seen to threaten the status quo and question internal power structures, rituals, and established ways of doing things.

In my experience, what seems to have helped to overcome this resistance has been to create a demonstration effect. As a colleague of mine recently reflected: although the theory of change is not an officially endorsed tool within IKEA, it has become a recognized practice as least within sustainability. This seems to have happened with theories of change being developed in isolated parts of the organization, and it has spread to other parts. As it has spread, more and more people have become exposed to it, resulting in increasing recognition of its value and usefulness within the organization.

There are also two tactics to consider as to the use and framing of a theory of change approach in a corporate context.

First, there is a need to make the theory of change tool as accessible as possible by being pragmatic about its use. This may involve working with a simple linear structure, without necessarily—at least in the first instance—specifying assumptions, risks, alternative pathways, narratives, and associated targets and key performance indicators, as is commonly recommended (for example, in the international development context). The aim is to encourage the use of the tool and not be formalistic about it, under the assumption that both the structure and creation process are value-adds in themselves.

In doing this, it is important to recognize that it is unlikely that the theory of change approach will ever be the main strategic tool used within a company. Instead, applying a theory of change approach could usefully be considered as an input or part of a company's prevailing strategic tools. Such integration is quite possible, given that the theory of change approach is not as foreign to the business world as it may seem. The basic output-outcome-impact logic of the theory of change approach is similar to established corporate strategic tools, such as balanced scorecard and objectives and key results. Additionally, since a theory change is usually developed through an inclusive process, it complements nicely the often top-down approach to traditional corporate strategic management (Reeves et al., 2018).

Second, as indicated above, to manage the evolving complexity of corporate sustainability, use of the theory of change approach should most meaningfully be seen as part of embarking on a learning journey. A learning organization is not based on planning but on feedback, which is an advantage in a complex and changing world that is difficult to understand and predict (Dahler-Larsen, 2011). This also aligns well with the fundamental characteristics of the theory of change approach, which is about managing complexity, identifying change, building on previous experiences, and encouraging open reflection and participation.

Sceptics may wonder if these two tactics risk hollowing out the theory of change concept, making it less 'rigorous'. What happens if we end up with a quickly assembled set of 'boxes' with unsubstantiated hypotheses about how they are connected, lacking explicit assumptions, risks, and alternative pathways? Or what if the theory of change does not address the tension between business and social objectives, is never finished, is not translated into action, or not owned beyond the engaged group of people that was involved in its development?

However, these are challenges that characterize all strategic exercises. In my work, a theory of change approach has clearly shown through both its structure and its creation process to help corporate employees to reflect on and create meaning and more clarity on prioritized topics related to the intersection between business and society. It is also clear that the right preconditions need to be in place for a theory of change to be put into practice, such as a clear change mandate, involvement of the key people, management support, complementary actions, and so on. Sometimes a theory of change approach may contribute to setting things in motion, sometimes not. But it is worth a try.

In conclusion, the theory of change approach is a powerful yet simple tool that helps create clarity and participation in the private sector, particularly in the intersection between society and commercial ambitions. As experts, we should promote its use by making it as simple and learning oriented as possible and not be held back by concerns over loss of rigour. In that way, we can truly contribute to making business more sustainable.

References

Bocken, N. M., Short, S. W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production*, *65*, 42–56. https://doi.org/10.1016/j.jclepro.2013.11.039

Dahler-Larsen, P. (2011). *The evaluation society*. Stanford University Press.

European Commission. (2021). *Social economy action plan—Employment, social affairs & inclusion*. Retrieved from https://ec.europa.eu/social/main.jsp?catId=1537&langId =en

European Venture Philanthropy Association. (2020). *What is venture philanthropy?* Retrieved from https://evpa.eu.com/about-us/what-is-venture-philanthropy

KPMG. (2020). *The time has come: The KPMG survey of sustainability reporting 2020*. KPMG Impact. Retrieved from https://assets.kpmg/content/dam/kpmg/xx/pdf/2020 /11/the-time-has-come.pdf

Meyer, C. (2018). *Prosperity: Better business makes the greater good*. Oxford University Press.

Murray, S. (2021). Measuring what matters: The scramble to set standards for sustainable business. *Financial Times*, May 14. www.ft.com/content/92915630-c110-4364-86ee -0f6f018cba90

Porter, M. E., & Kramer, M. R. (2006). The link between competitive advantage and corporate social responsibility. *Harvard Business Review*, *84*, 78–92.

Reeves, M., Legrand, J., & Fuller, J. (2018). *Your strategy process needs a strategy*. BCG Henderson Institute. Retrieved from www.bcg.com/publications/2018/your-strategy -process-needs-a-strategy

Rønholt Albertsen, R. (2021). *Are we asking the wrong questions in corporate social responsibility (CSR) research?* The Business of Society, Copenhagen Business School. Retrieved from www.bos-cbscsr.dk/2021/10/07/wrong-questions-in-csr -research/

Scott, M. (2021). *The 2021 global 100: How the world's most sustainable companies outperform*. Corporate Knights. Retrieved from www.corporateknights.com/ leadership/2021-global-100-progress-report/

WBCSD (World Business Council for Sustainable Development). (2013). *Measuring socio-economic impact: A guide for business*. World Business Council for Sustainable Development. Retrieved from www.wbcsd.org/contentwbc/download/2825/35536/1

World Economic Forum. (2019). *Reshaping global value: Technology, climate, trade—Global value chains under pressure*. World Economic Forum and United Nations Development Programme. Retrieved from www3.weforum.org/docs/WEF _Reshaping_Global_Value_Report.pdf

7 Upcycling theory of change for impact investment and early stage ventures

Penny Hawkins and Zazie Tolmer

Introduction

Impact investors and the initiatives they support aim to achieve social and/or environmental impacts alongside financial returns. Therefore, any related impact measurement framework needs to include a combination of financial data and impact data. This layering of evidence is designed to assist with understanding the causal connections between social or environmental changes and any potentially wider systemic effects. It also aims to better balance the dual priorities of achieving both financial and social or environmental impact.

The initiative we worked with was an impact investment fund that focused on improving worker well-being through making global supply chains more transparent and traceable using innovative technology. In using theory of change techniques in this market-based setting, we found we had to move beyond a conventional programmatic approach, and during the process we learnt different ways to articulate causal pathways to impact at the fund, venture, and market levels. The ventures we worked with are closer to the action than their investors and therefore tended to be clearer about their impact and how they expect changes to happen. What they needed was a tool that helped them evolve in line with their impact goals and bring together impact, revenue, and scale rather than only focusing on people and planetary impacts (for example, sustainable development goals). The fund, on the other hand, being more distant from impact goals and also engaged with a diversity of ventures at different degrees of proximity to the impact, needed a tool that would help make sense of their portfolio of investments in relation to impact and financial return, and consider more deeply their own role as an impact investor and the added value beyond financial investment.

At the fund level—from learning to strategy

As theory of change practitioners, we used a collaborative approach and worked in parallel on the theory of change process with the impact investors and several of their investee impact innovators. Throughout this process, rationales surfaced at the investor level theory that could be tested at the investee level and likewise from the investee back to the portfolio level. In this way, a more nuanced

and specific understanding of the opportunities and types of support needed for achieving impact and financial returns was generated.

We developed these theories of change through workshops and interviews, listening to both the fund team and venture teams' explanations and understanding of their contributions to impact. This information was then drafted into theories of change that we tested and refined during follow-up workshops. When new information became available, either through data collection activities or when new solutions were being developed or adapted, the theories of change were revisited.

The relationship between impact investors and innovators appeared to be well suited to this approach. Both parties' quest to identify solutions that have an impact and are profitable at scale drives a shared learning mindset open to inquiry and course correction while seeking results. This mindset fuelled opportunities for sharing what has been learned, as well as insights that grounded the theories of change with real and diverse perspectives at both the investor and innovator levels.

Throughout this process, different elements of the fund theory of change would come in and out of focus. We did not hold ourselves to comprehensively cover all elements of a theory of change, nor did we stick to a standardised process. At each theory of change engagement with the fund team, we would focus on the current learning and how this affected their theory of change.

For example, the initial version of a theory of change developed in the fund's first year of operation drew on the fund impact thesis and its first few investments. The impact focus was clear: the alignment of impact pathways to investments in different types of innovative solutions, along with some initial assumptions. The second iteration drew on theories of change developed with a few of the ventures and a round of interviews with the fund team to review the initial theory of change. This second version focused on the attributes of promising ventures and the ways in which they were attracting customers for their products and services. A subsequent third version drew on exploratory interviews with all ventures. In this latest version, there is a stronger emphasis on the fund's role in supporting ventures beyond the invested funds and on learning gained from the ventures' effectiveness in scaling their innovations. These iterations took place over several years as part of a developmental evaluation process aligned with the development of the fund, investees, and a growing number of investments over time.

At the venture level—combining impact, revenue, and scale

There are several ways in which impact, financial, and scaling strategies might combine in a theory of change. At the venture level, we found that organising the causal hypothesis into three nested theories of change helped to disentangle these causal pathways and to surface how they interact. Figure 7.1 displays

one option that proposes a theory of change model built on three nested and dependent theories of change in a three-tiered architecture for impact investors and innovators, comprising: the innovation theory of change, the core theory of change at the market level, and the expected effects as a result of scaling (scale theory of change).

Our description of the nested conceptual model in Figure 7.1 can be explained through a real example. Quizrr (2023) provides a 'digital training platform that enables de-risking in global supply chains by educating workforces from the bottom up on business-critical topics to drive behavioural change. Our platform enables you to track progress, get actionable insights, and address pain points proactively with your business partners'.

The venture theory of change focuses on the causal relationship between the innovation and the impact goal. This relationship and its complexity varies in accordance with the type of innovation; and so, rather than focus on its adoption, it articulates how the innovation leads to impact. In the Quizrr case, social practice theory (Shove & Pantzar, 2012) is applied to a suite of digital training packages to create the capabilities and conditions for generative dialogue between employees and managers on worker rights and working conditions in a variety of settings and contexts.

The core theory of change at the market level focuses on the venture's pathways to market. It identifies customers, explains why and how these clients purchase and use the innovation, and proposes how appropriate and continued use can be sustained. For Quizrr, this involves building their impact credibility with potential clients, ensuring their products are affordable, and developing a dashboard to support decisions and subsequent behaviour change. The dashboard makes a client's progress visible, helps in identifying pain points in their management systems, and highlights successes such as improved social audit scores.

The scaling theory of change links the two nested layers and unpacks the causal pathways between the innovation, its intended impacts, and financial returns. It considers how to scale the adoption of the innovation across the market—given the relevant incentives, barriers, and flows in the system—as well as the likely requirements, which might include the business growth capacity of the innovator, sufficient investment finance, or new business partnerships that extend their scope for further scale. The Quizrr scaling theory of change includes three strategies: (1) prove their business and impact case to potential clients by evidencing the innovation and market theory of change; (2) deepen their impact credentials by working in collaboration with the International Labour Organisation and others, in the development of their innovation; and (3) collaborate with complementary businesses that can help strengthen their business model.

The nested theory of change model also creates a useful framework in which potential impact, revenue, and scale measures can be identified. The theories of change can guide the development of a menu of measures, with each level bringing a new perspective on the expected effects of the impact investment.
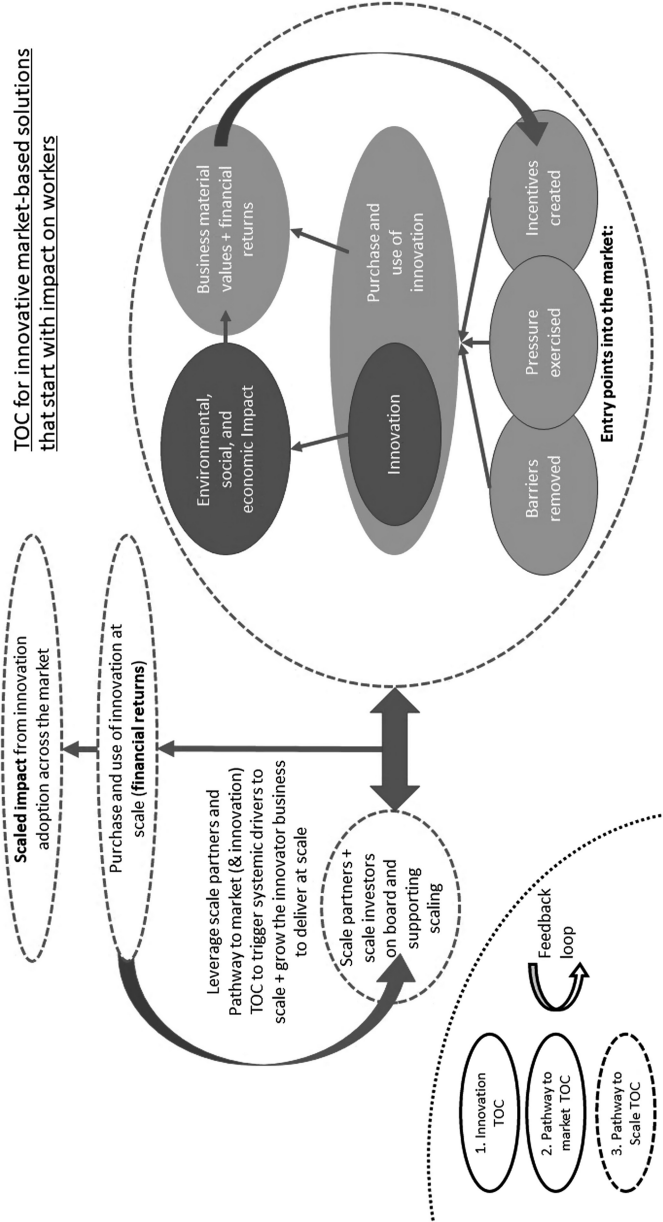
Figure 7.1 A theory of change for a venture combining impact, revenue, and scale

Once a menu of impact measures exists, the nested theories of change enable the selection of critical impact measures that, in combination, provide a focus for gathering evidence across impact, revenue, and scale.

At the venture level—creating space for iteration and improvement

For a few of the ventures, we were able to help shape theory of change archetypes that were useful in delineating the boundaries for the impact solution or investment and the broad causal pathways of change, whilst maintaining space for ongoing experimentation, learning, and expected and unexpected shifts.

We developed a theory of change for one venture that proposed a causal hypothesis of how their innovative service created impacts at scale using a structure based on a stakeholder analysis. This was developed through a series of dedicated online meetings with two members from the venture. We used these inputs to produce a version that organised the information into the main causal pathways, shifting away from the detail to the critical causal elements. After quickly testing this more focused version, the venture asked a graphic designer to make it look 'pretty'; this version was then tested with the rest of the venture leadership and staff. This enabled the venture to explain succinctly and clearly what it does, how it drives change, and how it eventually achieves intended impacts. They used this theory of change when pitching for further investment or to bring in major clients. They also used it when developing new services to confirm that the new service aligns with their theory of change and to understand what the new service might contribute to their intended impact.

Reflection

In using a theory of change approach in this market-based, impact investment context, we encountered some of the same limitations to a theory of change used in other settings—for example, using a conceptual model to capture 'on the ground' or in this case 'in the market' realities. Because there were no existing theory of change archetypes we could use, we had to be exploratory and at times creative to adapt our programmatic theory of change practice to the vicissitudes of continuous change with early stage ventures working on innovative solutions within a market context.

The theory of change tool has been strategically useful to ventures in articulating, testing, improving, and communicating their approach to impact in a market context. The nested model also helps with thinking about how to scale. For the impact investment fund, the theory of change has been more useful in articulating their role and in making sense of impact pathways and learning from venture experiences.

In the absence of an alternative, our experience with using the theory of change for this initiative found that it can add value, for both an impact investment

fund and the early stage ventures supported, to explore the causal relationships between the innovation, the market, and scaled impact and revenue.

References

Quizrr. (2023). Impact pathway. www.quizrr.se

Shove, E., & Pantzar, M. W. (2012). *The dynamics of social practice—Everyday life and how it changes*. Sage Publications.

8 Strategy-level theories of change require a focus on systems change

An actor-based approach can help

Andrew Koleros

Philanthropic foundations have historically played an important role in funding programmes that aim to transform the lives of the poor and other marginalized groups (OECD, 2018). But despite the vast number of social and economic reform programmes that foundations have created and supported over the past several decades, these efforts are often not leading to transformational changes in the lives of many of the world's populations most in need.

This has driven the philanthropic sector toward an improved recognition and understanding of the complexity of the problems that the sector is trying to address (Walker, 2017). It is now widely acknowledged that these 'complex problems', such as housing, unemployment, and health inequities, remain persistent due to a wider constellation of constraints that 'hold these problems in place' (Head, 2018; Kania et al., 2018, p. 3). These constraints range from government policies to societal and cultural norms, institutional arrangements, market dynamics and incentives, and power imbalances (Meadows, 2015).

Many leading philanthropic foundations now recognize that their conventional approaches of supporting stand-alone grants or siloed programmes are not enough. Without tackling this myriad of constraints collectively, the complex problems that the poor and marginalized face are likely to remain in place despite the sector's best efforts. This has led many foundations to now focus their strategies on achieving 'systems change'. It should be noted that this is a shift that many other actors in the development ecosystem have taken long before, with system approaches documented to overcome both public- and private-sector challenges (Beinhocker, 2006; OECD, 2017). Thus, multiple types of guidance and resources around systems change exist, with different groups defining 'systems change' through their respective sectoral lenses.

Kania et al. (2018) credit the Social Innovation Group in Canada with a definition of systems change that fits well with the ambitions of the philanthropic sector: 'systems change means shifting the conditions that are holding a problem in place'. They go on to suggest that foundations that want to embrace a systems change approach need to move beyond typical grant programmes and first take time to deeply understand the wider system that surrounds the complex problems they are trying to tackle.

By identifying all the actors that touch the issue they seek to address, foundations can identify a diverse set of organizations and individuals already working on the various constraints associated with complex problems. The role of the foundation is then to work to help these groups join forces to achieve their common goals. This collective effort thus represents a portfolio of investments within an overall foundation strategy aimed at systems change.

Although the approach resonates with many foundations at the conceptual level, operationally, the shift from conventional grant making to designing intentional strategies aimed at driving systems change is not straightforward. It requires foundations to adopt new ways of thinking—about how they understand the problems they are trying to address, as well as the ways in which their support can address them. For many, this is unchartered territory, requiring new approaches to assist them in this endeavor.

This essay makes the case that an actor-based approach to theory of change development can assist foundations in operationally making this shift. While actor-based theories of change have been used successfully to unpack complexity within international development programmes, they have not yet been widely adopted by the philanthropic sector for similar purposes. But given the systems focus of many foundation strategies, the approach has much to offer to philanthropy.

To this end, I first present the concept of an actor-based approach to theory of change development and work through an illustrative example to demonstrate its suitability for supporting foundations to take a systems lens. This is followed by a discussion on how this approach can also assist foundations in visually depicting and describing their strategy at different levels for different audiences. I conclude with a brief discussion on the implication of this approach for evaluators.

An actor-based approach can ensure that a strategy focuses on systems change

Among programme theorists designing theories of change for interventions in complex systems, there has been increasing attention given to the advantages of an actor-based approach that unpacks some of the complexity of international development interventions (Koleros & Mayne, 2019; Koleros et al., 2020). This section describes how an actor-based approach to theory of change development can similarly support foundations to design strategies through a systems lens using the example of a foundation strategy around youth's participation in the workforce.

An actor-based approach to theory of change development is rooted in complexity theory. That is, it takes the perspective that a complex problem can be understood as the system-level behavior that manifests as a result of the pattern of actions and interactions among individual and autonomous system actors. These actions and interactions occur between system actors and in response to

their environment (including their wider contextual conditions). Put another way, the different ways in which various actors within the system behave explain why a complex problem exists. This perspective is consistent with Kania et al.'s framing around the need to understand and identify the multiple constraints that 'hold a problem in place'.

From this theoretical perspective, an actor-based approach starts off with a systems-mapping exercise to understand what is holding a current problem in place. In this exercise, the constellation of actors surrounding a problem is identified, along with the current patterns of actions and interactions of these system actors—with each other and their wider environment.

For example, consider the case of low levels of young women entering the workforce. This can be understood as a system-level behavior that manifests as a result of the patterns of actions and interactions among different system actors that influence young women's employment outcomes within a given context. These include employers' practices around youth workforce recruitment, the interactions between employers and training service providers to ensure that skills training programmes sufficiently meet the market's needs, governmental policies and regulations that influence youth job creation, and the cultural and societal norms held by community members around the role of women in the workforce (see Figure 8.1).

Through this actor-based lens, addressing a complex problem requires shifting the actions and interactions among system actors over time from their *current state* (which explains the current state system behavior—or 'problem') to a *future state*, where the actions and interactions among system actors create a new system behavior.

From the shared understanding of the 'current state', a vision for the 'future state' system can be developed that describes what the system would look like when the problem no longer exists. Put another way, a vision for how the actions and interactions of system actors would need to change to generate the desired 'future state' system behavior.

This responds to Kania et al.'s guidance for foundations focusing on systems change to identify all the actors that touch the issue they seek to address, as well as the diverse set of organizations and individuals already working on the various constraints associated with the problem. In this way, the future state systems map can assist in identifying a set of potential grantees. Each grantee would be working to change the actions and interactions among different system actors to collectively address the wider system-level problem.

Taking the example from above, in a 'future state' where the problem of youth participation in the workforce is addressed, we would expect to see a number of actor-level changes, including changes in training programmes that better meet industry needs (as a result of improved interactions with potential employers), changes in youth workforce recruitment and development practices among potential employers, and shifts in mindsets among community leaders
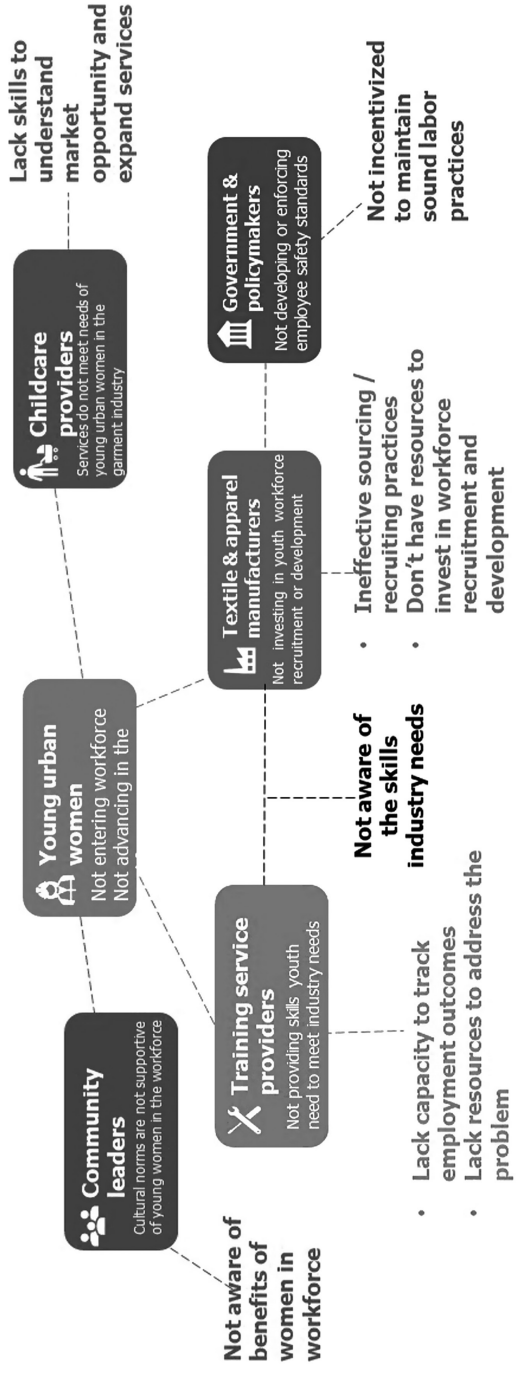
Figure 8.1 Example of a 'current state' actor-based system map around youth workforce participation

and other influencers around the role of young women in the workforce (Figure 8.2). This provides a basis for foundations to begin thinking about a series of grants to different individuals and organizations who would be able to shift the current practices among these different system actors within an overall portfolio.

The theory of change articulated through this process is essentially the posited pathways from foundation investments made to grantees in the current state to the visioned future state of the system as a result of grantee actions. The actor-based theory of change has two main components: (1) the theory of change around how individual grantee activities are meant to lead to changes in individual system actors (actor-level theories of change) and (2) the theory of change around how activating multiple actor-level changes is meant to lead to systemwide changes (changes in system conditions) and an overall change in system behavior (that is, addressing the complex problem).

The first component is operationalized through a series of actor-level theories of change that depict the detailed steps in the causal impact pathways from a given intervention to an actor-level change in practice (and the causal assumptions underlying these pathways). In this sense, there is not just one overall 'actor-level theory of change' but a series of nested theories of change for each specific system actor.

For instance, an actor-level theory of change to shift the practices of government and policymakers might be developed for a grant to an advocacy organization aimed at driving a policy change. Another might depict the actor-level pathway that grants to grassroots organizations take to conduct community organizing and civic education activities that address narrative change and shifts in mindsets. Another actor-level theory of change would be developed to demonstrate the pathway from a grant to a business development service provider to changes in the courses that training service providers provide so they are better targeted to meet industry needs (Figure 8.3).

The second component uses the future state system map to articulate a system-level theory of change that identifies how activating multiple actor-level changes (via actor-level theories of change) will lead to overall systemwide changes and contribute to changes in overall system behavior. In identifying possible systemwide changes, Kania et al. (2018, p. 4), drawing from the work of system thinkers including Jay Forrester, Donella Meadows, and others, posit six conditions of systems change articulated at three levels: (1) structural or explicit changes (changes in policies, practices, and resource flows), (2) relational or semi-explicit changes (changes in relationships and connections and changes in power dynamics), and (3) transformational or implicit changes (changes in mental models) (Figure 8.4).

Within the future-state actor-based system map, these six conditions of system change can be understood as either actor-level changes (for example, changes in practices or policies) or changes in the interactions of system actors with each other and their environment (for example, changes in resource flows,
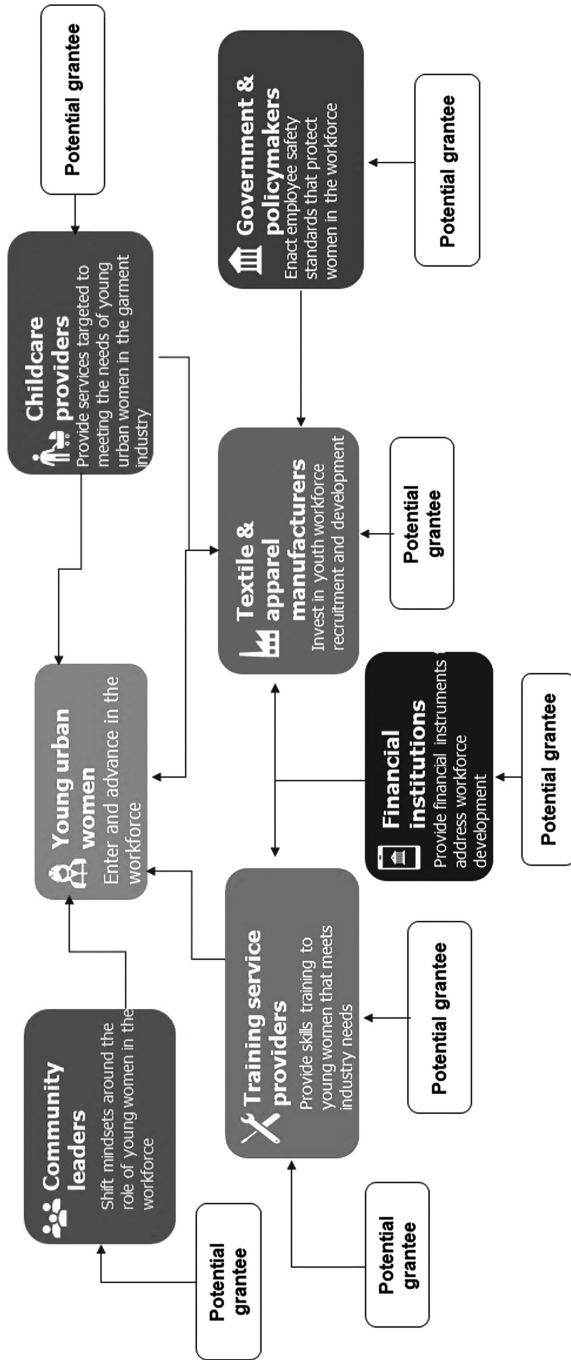
Figure 8.2 Vision of a 'future state' actor-based system map addressing the problem of youth workforce participation
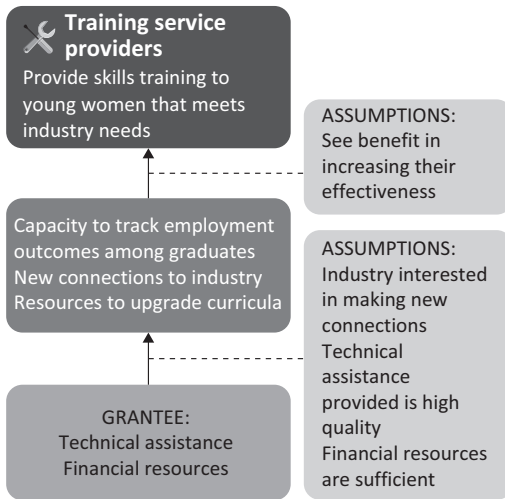
Figure 8.3 Example of an actor-level theory of change for training service providers
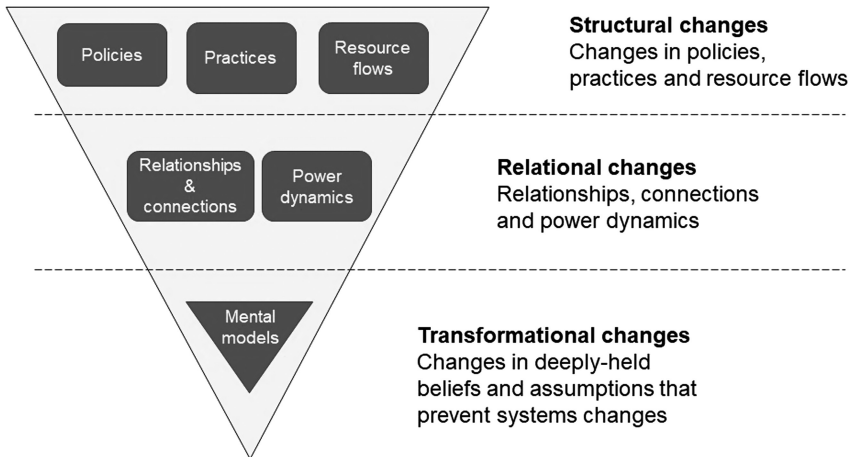
Figure 8.4 Six conditions of systems change

Source: Kania et al. (2018)

relationships, power dynamics, and others). Figure 8.5 identifies these six conditions within the future-state, actor-based system map using the youth workforce participation example. When these system changes occur, resulting in new patterns of actions and interactions between system actors, then changes
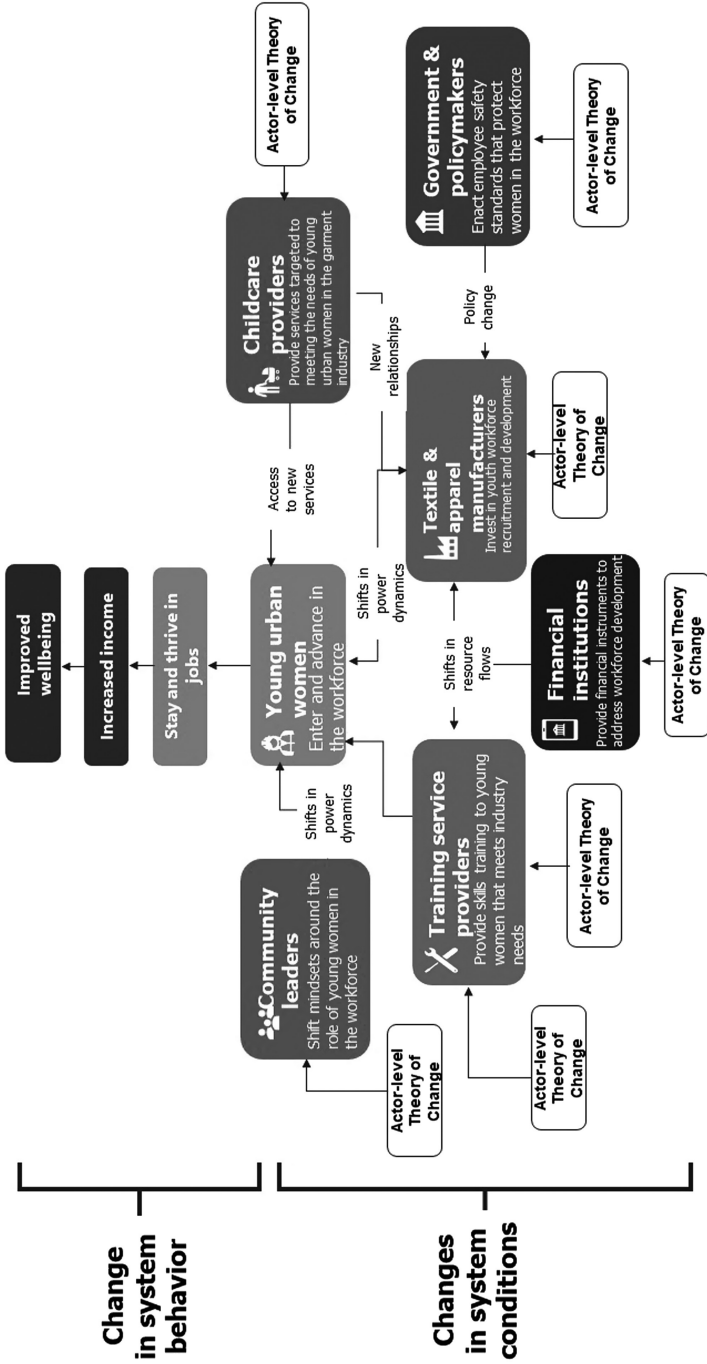
Figure 8.5 Actor-based systemwide theory of change for addressing workforce participation among young women

in overall system behavior are expected to occur. In the case of this example, we'd expect to observe higher employment rates among young women, leading to increased income and ultimately improved well-being at the individual, household, and community levels.

Using an actor-based approach to generate a theory of change that meets multiple uses for multiple users

In addition to assisting foundations in focusing on systems change, taking an actor-based approach to theory of change development can also support foundations in depicting and describing their strategy in different ways for different audiences or 'users' of the strategy. For instance, actor-level theories of change that articulate the detailed steps in the causal pathway from grantee activities to the intended changes in system actors may not only help grantees to understand how their work connects to the bigger picture, but they may also support foundation teams to measure over time the degree to which change is happening as expected in order to drive course correction and implementation improvements.

On the other hand, zooming out to the overall systems level, the systemwide theory of change allows the depiction of the strategy at the broad conceptual level needed by foundation leadership, thereby providing the overarching, high-level changes occurring across the portfolio without getting into the detailed or contextual information that would be needed at a grantee level. In between these two extremes, managers can use both actor-level theories and the wider systems-level theory of change to communicate how each individual grant is intended to bring about a change in a system condition and also how the collective efforts of grantees are intended to work together to address the myriad of constraints that hold the problem in place.

Implications for evaluators

As this essay has outlined, adopting an actor-based approach to theory of change development is a promising approach to assist foundations in shifting from conventional grant programmes to focus on system change. But foundation teams will need support to make the shift. Evaluators can play a strong role in this process, but will require a shift in the ways that evaluators engage with foundations as well.

In traditional philanthropy, evaluators are typically engaged once a strategy has already been developed—sometimes early in implementation to develop a monitoring, evaluation, and learning system for a strategy, or much later to conduct an evaluation. As many other essays in this book have described, a theory of change is a key tool in the evaluator's toolkit for these endeavors. However, assisting foundations to make the shift to strategies designed around systems

change requires evaluators to consider the theory of change not only as an evaluative tool but also as a programme design tool.

This implies that evaluators need to be engaged earlier and integrated into the actual strategy design process. In this capacity, evaluators need to adopt a design mindset as well as an evaluation mindset—seeing themselves as facilitators and thought partners to foundation teams, while infusing evaluative thinking into the strategy design process. In this regard, evaluators must also embrace the uncertainty of engaging in systems change efforts and help foundation teams to shift from traditional approaches of individual grant monitoring towards an appreciation of collective efforts and the need to generate learning to adapt and improve.

References

Beinhocker, E. D. (2006). *The origin of wealth: Evolution, complexity, and the radical remaking of economics*. Harvard Business School Press.

Head, J. (2018). Wicked problems in public policy. *Public Policy*. Retrieved from www.researchgate.net/publication/43502862_Wicked_Problems_in_Public_Policy

Kania, J., Kramer, M., & Senge, P. (2018). The water of systems change. Retrieved from www.fsg.org/resource/water_of_systems_change/

Koleros, A., & Mayne, J. (2019). Using actor-based theories of change to conduct robust evaluation in complex settings. *Canadian Journal of Programme Evaluation*, *33*(3), 292–315.

Koleros, A., Mulkerne, S., Oldenbeuving, M., & Stein, D. (2020). The actor-based change (ABC) framework: A pragmatic approach to program theory in complex systems. *American Journal of Evaluation*, *41*(1), 34–53.

Meadows, D. H. (2015). *Thinking in systems*. Chelsea Green Publishing.

OECD. (2017). *Systems approaches to public sector challenges: Working with change*. OECD Publishing. https://doi.org/10.1787/9789264279865-en

OECD. (2018). *Private philanthropy for development: The development dimension*. OECD Publishing. https://doi.org/10.1787/9789264085190-en

Walker, J. C. (2017). Solving the world's biggest problems: Better philanthropy through systems change. *Stanford Social Innovation Review.* https://doi.org/10.48558/743S-NV97.

9 Evaluating the use of artificial intelligence and big data in policy making

Unpacking black boxes and testing white boxes

Frans L. Leeuw

Background

Given the widespread diffusion of artificial intelligence/big data (AI/BD),[1] — recently conceptualized as 'the newest system technology' and compared in magnitude with the introduction of the steam engine, electricity, and the computer (WRR, 2021)—it is important that evaluators address the question of how to evaluate the expected, unexpected, and adverse effects of the use of AI/BD when designing, developing, and implementing interventions (of any kind): policies, programmes, regulation, therapies, drugs, and others.

Indicators of the societal role AI and BD play are manifold. They include mobile health and the Quantified Self movement, legal scholars and practitioners using machine learning in analyzing and drafting legal texts, law enforcement agencies predicting crime rates and patterns, medical professionals diagnosing and developing therapies and drugs, and policymakers designing and implementing programmes and policies (Dwivedi et al., 2021; Leeuw, 2021; Rajkomar et al., 2019; York & Bamberger, 2020; Zuiderwijk et al., 2021). Added to this are several artificial intelligence chatbots like ChatGPT by big tech companies that have been recently introduced.

At the same time, it is known that there are various problematic, complicated, and probably adverse effects of living in a 'society of algorithms' (Burrel & Fourcade, 2021). Examples are the bias problem (data collection only or largely includes subjects using apps, smartphones, or desktops), the legacy problem (working with 'old', sometimes biased data), the big data hubris problem (the belief that BD is a substitute for everything else, making theories and causal analysis obsolete), validity issues (Lazer et al., 2021), and the lack of transparency and explainability of what is happening when AI/BD are applied in developing and implementing interventions. Often this lack of transparency is related to the *black-box problem of AI/BD*—the focus of this essay.

Evaluating black boxes is familiar territory for evaluators; they attempt to uncover and test assumptions about contexts, mechanisms, and outcomes of activities (Astbury & Leeuw, 2010; Leeuw, 2020; Lemire et al., 2020; Nielsen

et al., 2021; Pawson, 2013). However, rarely have realist, theory-driven evaluators applied this approach to *black boxes when BD/AI are involved* in domains including health, law enforcement, labor, and education, among others.[2] This essay attempts to outline an approach to do that.

Unpacking black boxes of AI/BD

The *AI black box*[3] refers to the phenomenon that with most AI and AI-based tools *one does not know how they do what they do*. This problem has several characteristics:

- Although AI-based tools are often clear about the information on the input (the question or data the AI tool starts with), and that also usually applies to the output (the answer), it often remains unclear how the input is turned into the output and what the respective roles of algorithms and (big) data are. As Price and Rai (2021) mention, 'even when human field experts are given full access to the learning algorithm, training data, training process, and resulting model, the models can be difficult to parse because they are often complex and nonintuitive' (p. 779). They distinguish between two related layers of what is called *opacity:* 'the opacity of the system being studied, and the opacity of the research tool (machine learning) being deployed to study it' (p. 778).
- Another characteristic next to *opacity* and the related lack of *explainability of* AI/BD is *plasticity*, which means that the algorithms change in response to new data. Price (2018) mentions, for the field of medicine, that 'this form of frequent updating is relatively common in software but relatively rare in the context of other medical interventions, such as drugs, that are identified, verified, and then used for years, decades, or even centuries' (p. 1).
- The third characteristic is that the *different actors* engaged in using AI/BD in organizations like hospitals, governments, and companies have often *different levels of practical expertise and experience*. Some clinicians, policymakers, administrators, or managers are more 'into' AI and BD than others. Probably this also applies to auditors, inspectors, and other oversight officials. Ranerup and Henriksen (2020, p. 1) studied the introduction of robotic process automation (RPA) into the world of (governmental) social services and what it did to the civil servants. Apart from positive effects, they find 'that a human–technology hybrid actor [RPA] redefines social assistance practices. Simplifications are needed to unpack the automated decision-making process because of the technological and theoretical complexities'.

Black boxes are not a 'given'; *they can be unpacked* and *made* into white boxes, that next need (external) validation (or testing). The question is: *Does working*

with AI/BD contribute to the effectiveness of interventions?* The next section offers a six-step approach to help find that out.

A six-step approach to unpack AI/BD black boxes into white boxes and test them

This section discusses several steps evaluators could follow when they start to think about unpacking black boxes in the world of artificial intelligence.

Step 1

The first step is to specify the *goals or the contributions to be achieved* when applying AI/BD in designing, developing, and implementing interventions (see Table 9.1).

Step 2

This step concerns the identification of assumptions that underlie the processes when interventions are designed, developed, and implemented using AI/BD. Assumptions are sometimes 'hidden' (Pawson, 2008), which means that they have to be articulated. Bennett Moses and Chan (2018) did that for predictive policing, Mitchell et al. (2021) for 'algorithmic fairness', Kempeneer (2021) for the 'big data state of mind', and Domingos (2015) for theory-families ('tribes') existing in AI.[4] This step tells us at a minimum that—contrary to the big data hubris-claim—theories are important, and they can and will differ (as do the criteria that can be used to 'judge' or 'test' them). *Put differently: AI/BD black*

Table 9.1 Examples of goals and contributions to be achieved

Example	Goal or contribution of using AI/BD
Digital alternative dispute resolution	An 'intelligent agent' is the AI application. These agents can be a tool for the adjudicator (reviewing all documents; researching similar cases) or they can be the arbitrator. The goal is to increase efficiency, trust, and effectiveness in and of the dispute solution.
High-performance medicine	Algorithms are used with the goal to detect (often together with human intelligence) pneumonia, do medical scans, carry out diagnoses of skin cancer, and check eye conditions (Topol, 2019).
Education	Computerized adaptive tests are a form of machine-learned assessment, used with the goal to optimize summative assessments in high-stakes selection processes.
Insurance	AI is used for more effective and efficient illness and disability claim prediction and for fraud detection.

boxes have to become white boxes. In order to do that, a **framework** is needed that *specifies dimensions of AI/BD and their use*. Based on earlier studies, Figure 9.1 is the visualization of such a framework. It combines the context-mechanism-outcome model from realist evaluations with components and characteristics of the data and algorithms and the types of machine learning applied.

The central piece in the framework is the *use* of AI/BD in decision making and its (assumed) contribution to the effectiveness of the interventions. A *core assumption* is that the more AI/BD plays a role in these processes, the larger the likelihood that the interventions will be (more) effective than when AI/BD does not play that role. This is *believed* to happen *because* the design, production, and implementation processes are:

- free from human failures, including fatigue, cognitive and computational restrictions, and personal biases
- always up to date (using diverse types of data, including real-time data)
- precise in their focus
- free from implementation failures
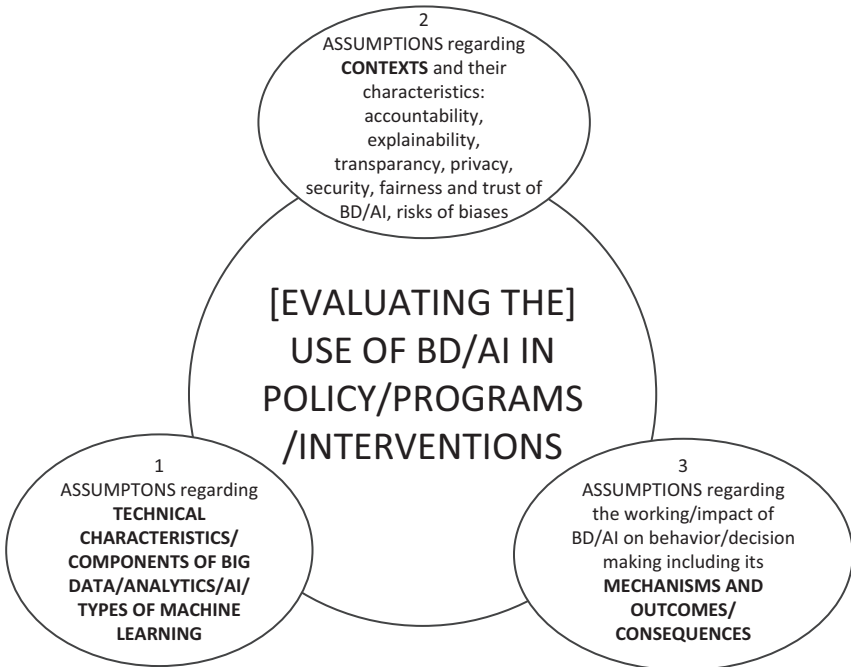
Figure 9.1 Framework for specifying dimensions of AI/BD and their use

Circle 1 includes assumptions on the *type and quality and relevance of data and its data-ecosystem*, including types of *machine learning* (like reinforcement learning, deep learning, and so on). Examples of these types of assumptions include:

1. If data are used, they will accurately reflect reality.
2. The future is like the past (or: the best predictor of the future is the past).
3. Data analytics does not unjustly discriminate in terms of gender (for example).

Circle 2 regards assumptions on the *(societal) contexts* in which AI/BD is used when interventions are designed, developed, and implemented. Pawson et al. (2005) suggest that contextual factors must be considered at four different levels: **individual capacities** of the key actors and stakeholders, such as interests and attitudes; **interpersonal relationships** required to support the intervention, such as lines of communication; **institutional settings** in which the intervention is implemented, such as the culture of organizations[5]; and the **wider (infra-) structural systems** of a society. An example is given by Price (2018), who distinguishes between working with AI in medicine 'in high-resource contexts, such as academic medical centers versus in low-resource settings such as community health centers or rural providers in less-developed countries'.

Examples of these types of assumptions include:

4. If the data point to implementing certain (policy) actions, then it is assumed that 'perfect implementation' will take place.
5. If the goal is to realize social acceptability (and impact), attention must be paid to transparency and accountability of the working processes, fairness, explainability, nonmaleficence, responsibility, security, privacy, reliability, and trust. This also requires putting human values at the core of AI-systems.
6. If this goal is to be achieved, organizations should implement general (information technology) controls to ensure that their IT systems are reliable and ethically sound.

Circle 3 looks 'inside' the operations when algorithms and big data are used and addresses assumptions underlying these operations. They include assumptions on the actors and their behavior that work with AI/BD and how they 'deal' with the challenges attached to issues like opacity and plasticity: Who are the actors? What are their perspectives on, knowledge of, and attitudes toward AI/BD? Which stakes are involved? The evaluator also has to look for indicators of impact/behavioral and social consequences, respective of the costs and benefits.[6]

Pedersen and Johansen (2020, p. 520 ff) introduced the concept of *behavioral AI* (BAI). Studying BAI is believed to open up the 'link between AI-behavior and AI-inference by describing how to study AI behavior'. Of particular importance in BAI are:

- The relation—similarities and differences—between human cognition and algorithmic processing
- The relation between human learning and algorithmic (machine) learning
- The process of inferring knowledge from data, thus arriving at valid and reliable judgments, made by an AI system compared to how humans make judgments.[7]

Examples of these types of assumptions include:

7. If one works with AI/algorithms, then (sometimes) it is assumed that the decision making will be far more efficient and fair than humans are ever capable of.
8. If there is eradication of the human factor in decision making, the one-sided focus on efficiency—and the use of computational analyses for control, surveillance, and prevention—could lead to a more critical attitude toward assumption number 7.
9. Algorithms are (always) neutral.
10. Algorithms (often) have a serious degree of plasticity, changing in response to new data (that is, frequent updating).

Step 3

The assumptions that have been *surfaced* are now a step closer to a white box, but that does not guarantee validity or truth; they need—as is always the case with (small-t and capital-T) theories—to be tested. Questions from a realist evaluator's perspective include:

- How relevant, valid, and reliable are these assumptions?
- Are they valid and reliable in a general way (that is, for 'all' BD/AI working processes) or only given certain contexts?
- Do they specify and articulate which mechanisms are involved in BD/AI operations in practice?
- Are the outcomes readily available or directly operational, measurable, and explainable?

Step 4

Assessing the validity of the (articulated) assumptions can be done *first* by using *existing empirical evidence* from (interdisciplinary) research in which *similar or look-alike AI/BD tools and cases* are investigated. These studies can be found in the social and behavioral sciences (like Bennett Moses & Chan, 2018), in behavioral computer sciences, and in computational social sciences and studies dealing with machine–human interactions (Lazer et al., 2021; Bowser et al.,

2021). They can help by transferring that knowledge to one or more look-alike AI/BD cases to help make predictions about the probable validity of the AI/BD used. Sometimes this approach is called 'subsuming interventions or cases under general theories' (Leeuw, 2012; Pawson 2002a, 2002b) or framed by Foy et al. (2011, p. 454) as 'generalization through theory'.

Step 5

Predicting the probable validity of a white box *based on existing research on look-alikes is oftentimes not enough*. New, *primary research is needed*. In the literature, a distinction is made between *in silico studies*—evaluating the algorithms and data operations as such (that is, operate within and between computers) and *behavioral evaluations*—evaluating the implementation and contributions of the algorithms and big data in practical, real-life situations. We do not discuss in silico evaluations here and only focus on the second 'type' of evaluations.

Focusing on the world of medicine, Price (2018) described three activities that need to be done in such an evaluation.

- **Step 5.1.** The first activity is 'ensuring that algorithms are developed according to well-vetted techniques and trained on high-quality data' (p. 2).
- **Step 5.2.** The second concerns reliability: 'demonstrating that an algorithm reliably finds patterns in data. This type of validation depends on what the algorithm is trying to do. Some algorithms are trained to measure what we already know about the world, just more quickly, cheaply, or accurately than current methods… Showing that this type of algorithm performs at the desired level is relatively straightforward… Other algorithms optimize based purely on patient data and self-feedback without developers providing a "correct" answer, such as an insulin pump programme that measures patient response to insulin and self-adjusts over time. This type of algorithm cannot be validated with test datasets' (p. 2).
- **Step 5.3.** The third activity 'applies to all sorts of black-box algorithms: they should be continuously validated by tracking successes and failures as they are actually implemented in health-care settings' (Price, 2018, p. 2). For performance one can also read: impact or effects of the AI-based intervention when dealing with patients/clients in real life. Park and Han (2018, pp. 806–807) add this: 'With a computerized decision-support system such as artificial intelligence, not only its technical analytic capability but also the way in which the computerized results are presented to, interpreted by, and acted on by human practitioners in the clinical workflow could affect the ultimate usefulness of the computerized algorithm'. They suggest to use randomized controlled trials to sort this out, but they are also open to other designs. Vijayakumar and Cheung (2021) add that checking replicability of

AI/machine learning-based results is strongly recommended. An application to the world outside medicine is presented by Choenni et al. (2021) for the field of using AI/BD in smart cities.[8]

Step 6

This step concerns the transfer of the findings to experts, other professionals, and society at large. The goal is to inform parties and society about the validity of the approach, which is intended to help *explain* how BD/AI has been applied in the process, show the transparency of that process, and increase its social acceptance.

Conclusions

This essay outlined the relevance of thinking in line with realist (theory-driven) evaluations to unpack and test AI/BD black boxes. It included a six-step approach. Because human–machine interaction is involved—together with a continuous flow of data, plasticity of algorithms, and different types of machine learning—this is not an easy task.

If the statement 'practice makes perfect' is correct, then that is the way to go. This should include learning from what is already happening in other worlds, like in medicine.

All this may and probably will help increase the relevance of evaluating AI/BD-driven interventions and policies and contribute to an effective, ethical, and socially acceptable 'Algorithmic Society'.

Notes

1 For readers not familiar with the concepts of big data, AI, and machine learning, see Janev (2020), www.linkedin.com/pulse/intelligent-things-its-all-machine-learning-roger-attick/ and www.zendesk.com/blog/machine-learning-and-deep-learning/.
2 See Bamberger (2016); York & Bamberger (2020); https://datapopalliance.org/lwl-27-the-role-of-big-data-and-ai-in-monitoring-and-evaluation-me/; and Rathinam, F. et al. (2020).
3 Sometimes one refers to 'alchemy' or 'black art' when characterizing AI black boxes (Campolo & Crawford, 2020, 7 ff).
4 Examples are the symbolists, the evolutionaries and the Bayesians; he described their characteristics, including assumptions they work with.
5 Sometimes reference is made to 'cultural metacognitions' that exist in organizations. They regard the knowledge of and control over thinking and learning activities in organizations, like the awareness of different contexts, analyzing them, and developing plans of actions for different cultural contexts.
6 One example is Ranerup and Henriksen (2020, p. 5) investigating 'Trelleborg, the first municipality in Sweden to use automated decision making for social assistance. The Trelleborg Model is a management model now used in many other municipalities in Sweden'. A second example is a study of AI adoption in public sector organizations, comparing three cases in three countries (van Noordt & Misuraca, 2022).

7 They add this: 'Behavioral Artificial Intelligence (BAI) would study the artificial inferences inherent in, and the manifested behavior of, artificial intelligent systems in the same way as the social sciences have studied human cognition, inference and behavior'.

8 Contribution analysis may also be an interesting approach to apply. The main reason is that AI/BD are not alone in making and implementing policy programmes/interventions; they always act in combination with human intelligence, experiences, prior individual knowledge, and so on. So, the focus of an empirical investigation would probably be most relevant if it tries to sort out what the contribution of AI (in interaction with humans) has been in developing and implementing programmes and interventions.

References

Astbury, B., & Leeuw, F. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation*, *31*(3), 363–381. https://doi.org/10.1177/1098214010371972

Bamberger, M. (2016). *Integrating big data into the monitoring and evaluation of development programmes*. Global Pulse.

Bennett Moses, L., & Chan, J. (2018). Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society*, *28*(7), 806–822.

Bowser, A., Carmona, A., & Fordyce, A. (2021). *Unpacking transparency to support ethical AI*. Science and Technology Innovation Program, Wilson Center.

Burrel, J., & Fourcade, M. (2021). The society of algorithms. *Annual Review of Sociology*, *47*, 23.1–23.25.

Campolo, A., & Crawford, K. (2020). Enchanted determinism: Power without responsibility in artificial intelligence. *Engaging Science, Technology, and Society*, *6*, 1–19. https://doi.org/10.17351/ests2020.277

Choenni, R., et al. (2021). Exploiting big data for smart government: Facing the challenges. In J. C. Augusto (Ed.), *Handbook of smart cities* (pp. 1–23). Springer. https://doi.org/10.1007/978-3-030-15145-4_82-1

Domingos, P. (2015). *The master algorithm*. Basic Books.

Dwivedi, Y., et al. (2021). Artificial intelligence: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice, and policy. *International Journal of Information Management*, *57*(April), 1–47. https://doi.org/10.1016/j.ijinfomgt.2019.08.002.

Foy, R., et al. (2011). The role of theory in research to develop and evaluate the implementation of patient safety practices. *BMJ Quality & Safety*, *20*(5), 453–459.

Janev, V. (2020). Ecosystem of big data. In V. Janev, D. Graux, H. Jabeen, & E. Sallinger (Eds.), *Knowledge graphs and big data processing* (pp. 3–19). Springer. doi.org/10.1007/978-3-030-53199-7_1

Kempeneer, S. (2021). A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation. *Government Information Quarterly*, *38*(3), 1–8. https://doi.org/10.1016/j.giq.2021.101578.

Lazer, D., et al. (2021). Meaningful measures of human society in the twenty-first century. *Nature*, 595, 189–196.

Leeuw, F. (2012). Linking theory-based evaluation and contribution analysis: Three problems and a few solutions. *Evaluation*, *18*(3), 348–363.

Leeuw, F. (2020). Program evaluation B: Evaluation, big data, and artificial intelligence: Two sides of one coin. In E. Vigoda-Gadot & D. R. Vashdi (Eds.), *Handbook of research methods in public administration, management and policy* (pp. 277–297). EE Publishers. https://doi.org/10.4337/9781789903485

Leeuw, F. (2021). Big data, artificial intelligence, and the future of evaluation. Background report to a presentation given at the Seminar of the Evaluation Network of DG Regional and Urban Policy, July 1, 2021.

Lemire, S., Kwako, A., Nielsen, S. B., Christie, C. A., Donaldson, S. I., & Leeuw, F. (2020). What is this thing called a mechanism? Findings from a review of realist evaluations. *Causal Mechanisms in Program Evaluation*, *2020*(167), 73–86.

Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, *8*, 141–163.

Nielsen, S., Lemire, S., & Tangsig, S. (2021). Unpacking context in realist evaluations: Findings from a comprehensive review. *Evaluation*, *28*(1), 91–112.

Park, S. H., & Han, K. (2018). Methodologic guide for evaluating clinical performance and effect of AI technology for medical diagnoses and prediction. *Radiology*, *286*(3), 800–809.

Pawson, R. (2002a). Evidence-based policy: The promise of 'realist synthesis'. *Evaluation*, *8*(3), 340–358.

Pawson, R. (2002b). Evidence and policy and naming and shaming. *Policy Studies*, *23*(3), 211–230. https://doi.org/10.1080/0144287022000045993

Pawson, R. (2008). Invisible mechanisms. *Evaluation Journal of Australasia*, *8*(2), 3–13. https://doi.org/10.1177/1035719X0800800202

Pawson, R. (2013). *The science of evaluation: A realist manifesto*. Sage.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review—A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy*, *10*(1), Suppl 1, 21–34. https://doi.org/10.1258/1355819054308530.

Pedersen, T., & Johansen, C. (2020). Behavioural artificial intelligence: An agenda for systematic empirical studies of artificial inference. *AI & Society*, *35*(3), 519–532. https://doi.org/10.1007/s00146-019-00928-5

Price, W. (2018). Big data and black-box medical algorithms. *Science Translational Medicine*, *10*(47). doi:10.1126/scitranslmed.aao5333.

Price, W. N., & Rai, A. K. (2021). Clearing opacity through machine learning. *Iowa Law Review*, *106*, 775–812.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347–1358.

Ranerup, A., & Henriksen, H. (2020). Digital discretion: Unpacking human and technological agency in automated decision making in Sweden's social services. *Social Science Computer Review*, *40*(2), 445–461. https://doi.org/10.1177/0894439320980434.

Rathinam, F., Khatua, S., Siddiqui, Z., Malik, M., Duggal, P., Watson, S., & Vollenweider, X. (2020). Using big data for evaluating development outcomes: A systematic map. CEDIL Methods Working Paper 2. Centre of Excellence for Development Impact and Learning.

Topol, E. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(January), 44–56.

Van Noordt, C., & Misuraca, G. (2022). Exploratory insights on artificial intelligence for government in Europe. *Social Science Computer Review*, *40*(2), 426–444. https://doi .org/10.1177/0894439320980449

Vijayakumar, R., & Cheung, M. (2021). Assessing replicability of machine learning results: An introduction to methods on predictive accuracy in social sciences. *Social Science Computer Review*, *39*(5), 768–801. https://doi.org/10.1177/0894439319888445

WRR. (2021). *Mission AI: The new system technology*. Netherlands Scientific Council for Government Policy.

York, P., & Bamberger, M. (2020). *Measuring results and impacts in an age of big data: The nexus of evaluation, analytics, and digital technology*. Rockefeller Foundation.

Zuiderwijk, A., Chen, Y., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and research agenda. *Government Information Quarterly*, *38*(3), 1–19.

10 Developing, representing, and using theories of change for interventions in complex systems

Patricia Rogers

What is meant by complexity and complex

The terms *complexity* and *complex* are used in many different ways in evaluation, as shown in a recent review of practice (Gates et al., 2021). A useful framing of complexity for evaluation was provided in a recent doctoral dissertation:

> 'Complexity' refers to an interdisciplinary understanding of reality as continuously forming deeply unpredictable patterns of behaviour and structure that inevitably emerge from interaction between diverse and interdependent components and cannot be reduced to or derived from the sum of these components. It is inherently dynamic. (Elliott, 2020, p. viii)

The Centre for Evaluation of Complexity Across the Nexus (CECAN), a UK-based research centre, also pointed to the inherent unpredictability of complex systems and identified three distinct features that contribute to this: (1) many diverse, interacting components; (2) nonlinear and nonproportional interactions between these components; and (3) in complex adaptive systems, adaptation or learning by the components in response to change (CECAN, 2020).

Nonlinear relationships are, of course, not only found in complex systems. But complex systems involve unpredictable nonlinear relationships—for example, involving thresholds or tipping points where things suddenly behave differently, and involving many, many variables.

Complexity should not be considered a fringe interest. The unexpected shifts caused by the COVID-19 pandemic and the extreme weather events caused by climate change have made the relevance of complexity more evident, but it also is an intrinsic part of many interventions. Many projects, programmes, and policies can be more usefully thought about in terms of interventions in complex systems (a framing suggested by Moore et al., 2019).

Complexity ideas are relevant everywhere (except perhaps in the most trivial of interventions), but they may be most useful, and easiest to get engagement with, when there is explicit intention to change a broader system or recognition of an individual or community agency—for example, leadership development programmes, participatory development projects, and individual-focused

programmes that work through supporting clients to identify and work towards specific goals.

Challenges for traditional approaches to theories of change

Changes are needed to how theories of change are developed, represented, and used to address the challenges faced when trying to use them in complex systems:

- The inherent unpredictability of complex systems that mean causal pathways cannot be confidently articulated in advance
- The numerous, diverse, and interacting components of complex systems that make them impossible to include in any representation
- The importance of agency, self-organisation, learning, and adaptation
- Nonlinear relationships that are hard to show in classic representations

Intended use of theories of change

Theories of change might at first seem inappropriate to use under conditions of ongoing uncertainty. Certainly, the rigid, linear version of theories of change, seen as a blueprint for implementation, is not appropriate for guiding and informing action during unpredictable change where adaptation is required repeatedly. Instead, the theory of change needs to be treated as a framework for helping people to think about the intervention and the systems in which it operates—not as a perfect representation of it that can be used for rule-based action and prediction.

Measures and indicators are an important part of using theories of change to gather and make sense of evidence, but moving to them too early carries a high risk of narrowing efforts to what can be readily measured rather than what really matters, with the common dysfunctional effects that follow of goal displacement and data corruption (for an overview of these issues, see Perrin, 1998). It will not be useful to develop monitoring and evaluation (M&E) systems based only on specific indicators from the initial theory of change. Even if the overall change theory remains intact, given ongoing adaptation, it is likely that there will be changes in terms of appropriate targets, and which specific activities, results, and contextual factors will be identified as most important for M&E. It might be more useful to develop rubrics, or global scales, that can include diverse evidence and be more readily updated over time through a systematic process. Accountability will need to be in terms of accountability for learning and appropriate adaptation rather than compliance with targets that are no longer relevant or beneficial.

Acknowledging the nonlinear relationships and variations among cases and over time, it will be important to develop M&E systems that are not simply univariate indicators and averages but that also provide analysis of patterns and

trajectories, including multivariate analysis and narratives and identification of important positive and negative feedback loops and tipping points.

Since interventions—and the systems in which they are implemented—are made of numerous, diverse, interacting components, theories of change need to be useful for the various actors implementing activities related to these components. They should therefore support planning, management, implementation, monitoring, and evaluation by the different actors involved.

Self-organisation and agency require adequate support for adaptation and learning. This requires a focus on the use of theories of change to support adaptation and learning as a primary use. This will require quite a shift in mindset for organisational stakeholders who use an agreed theory of change primarily or exclusively to generate and report performance information for accountability purposes and do not use it to support their own management and implementation. The theory of change needs to be a living document that is known by all those involved in implementation and used in review processes.

In his reflections on an evaluation of an integrated pest management programme, Douthwaite and co-authors concluded:

> Self-evaluation, and the learning it engenders, is necessary for successful project management in complex environments. (Douthwaite et al., 2003, p. 262)

Theories of change need to support double-loop learning (identifying incorrect assumptions and revising the theory of change) as well as single-loop learning (error and gap detection and correction).

An example of this broader type of use can be seen in the use of strategy testing by the Asia Foundation. This process begins with an initial theory of change, which is then reviewed in strategy testing exercises about every four months. After a discussion of the major developments since the previous review, the testing then moves to a series of questions about the existing strategy in terms of its various levels and components. The event ends with agreed revisions to the strategy and plans (Ladner, 2015).

Representations of theories of change

Instead of the usual 'intervention-centric' theory of change, which places the intervention as the focus, representation needs to show systems in which the intervention intends to cause a change, with deliberate decisions about where to draw the boundaries of the systems involved.

The overall format might need to be quite different—moving away from boxes and arrows with their implied precision and if-then logic. This theory of change might be expressed in terms of principles, which can be implemented in context-appropriate ways while still being capable of being evaluated. The
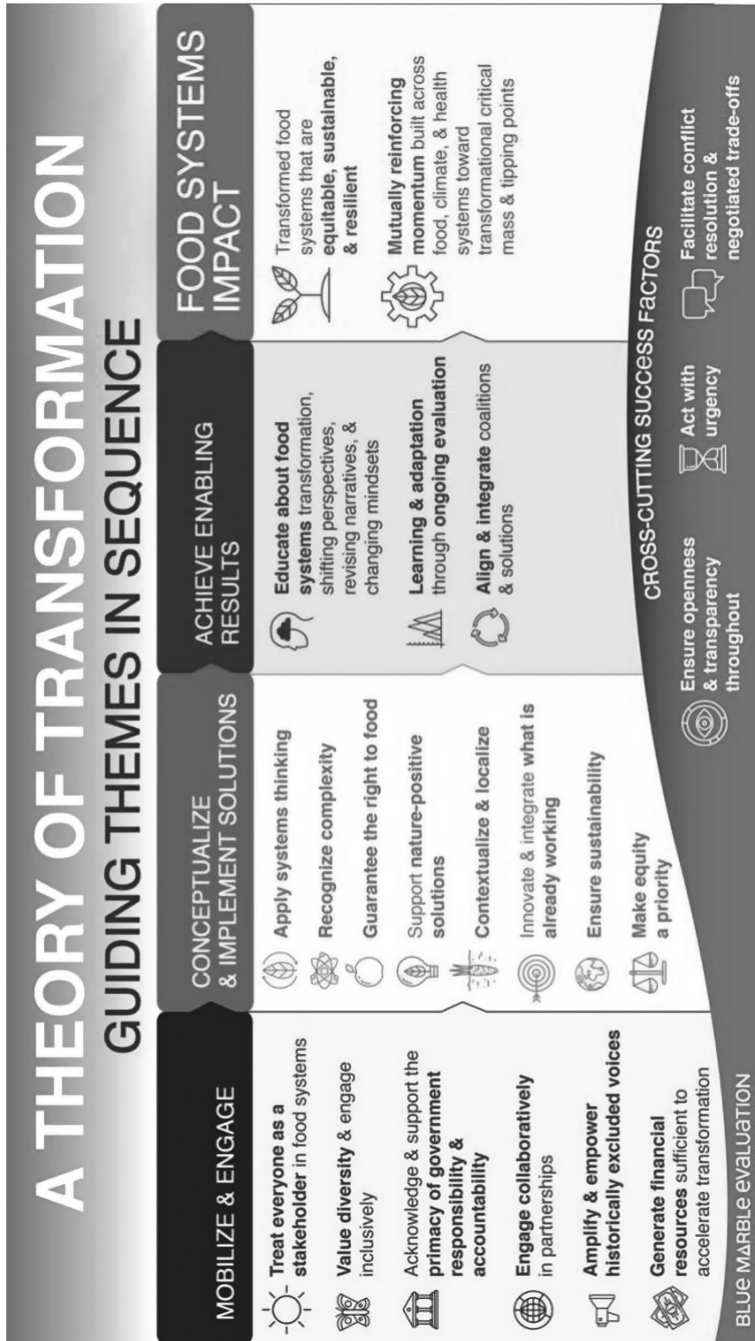
Figure 10.1 An example of a principles-based theory of change
Source: Patton (2020)

theory of change from the Global Alliance for Food Systems—a theory of transformation—provides a useful example of this—and also of how to turn its underlying principles into a framework for evaluation (see Figure 10.1).

When diagrams are used, they might be done in ways to make it clear that they are indicative heuristics: dotted lines to indicate causal relationships that are more speculative; multiple dotted lines to show alternative scenarios; and annotations about possible tipping points or thresholds, levers, and hubs.

Another major change in representing theories of change might be to move from a single diagram (with or without narrative) to multiple representations of theories of change. There might be different levels of focus—from macro- to meso- to micro-focus on specific causal links. There might be different causal pathways for different contexts, including a range of possible scenarios.

For example, nested theories of change have been developed for social accountability programmes, which support processes by which citizens and civil society organisations can hold the state and service providers accountable (Guerzovich et al., 2022); different combinations of these subtheories can be drawn on for particular contexts or combinations of contexts. The essays by Drew Koleros and Penny Hawkins and Zazie Tolmer in this chapter provide further examples of nested theories of change.

None of these representations of complexity-congruent theories of change can be readily fitted within the 'pipeline' format of logic models (inputs–processes–outputs–outcomes–impacts). Outcomes hierarchy formats will provide much more scope for showing other factors influencing results and feedback loops.

Processes for developing theories of change

The processes will need to be collaborative (to understand multiple perspectives and needs) and iterative (to incorporate changes in the situation and in the understanding of the situation). The essay by Hawkins and Tolmer provides a detailed example of this in impact investing.

An iterative approach between the specific and the general is likely to be helpful in building shared mental models grounded in shared or vicarious experience. Appreciative Inquiry or positive deviance approaches might be useful to identify what is working in specific examples, analyse why (including serendipitous juxtapositions of events, resources, and opportunities), and explore how this might be expanded.

Active engagement from diverse interests is needed in the processes of developing and revising the theory of change, including identifying potential nonlinear relationships and possible alternative scenarios. Engagement of diverse actors will also make it more likely that they will see the theory of change as useful and hence support the ongoing adaptation and learning required.

Table 10.1 Implications of features of complex systems for using, developing, and representing theories of change

Feature	Implications for how a theory of change is intended to be used	Implications for representing a theory of change	Implications for developing a theory of change
Uncertainty	Treat it as a useful heuristic not a blueprint	Make it clear in the diagram and narrative that the causal chains are not fixed or predictable but simplifications and probabilities Identify causal links with less uncertainty	Have an ongoing process of updating it as understanding improves, not expecting to have a final version or a complete understanding
Self-organisation and agency, adaptation and learning	Support self-organisation and agency by primarily focusing on supporting actors to learn and improve Support double-loop learning (revising assumptions) and single-loop learning (error detection and correction)	Develop representations that can be understood and used by different actors	Draw on quite different change theories around convening, knowledge exchange, co-generation, and creative emergence Engage local actors in the process as part of supporting self-organisation, local agency, and adaptation Continue to adapt the theory of change as priorities and possibilities change, without people feeling they are not making progress

(*Continued*)

Table 10.1 (Continued)

Feature	Implications for how a theory of change is intended to be used	Implications for representing a theory of change	Implications for developing a theory of change
Numerous, diverse, interacting components	Develop flexible, adaptable M&E systems Develop M&E systems that are about patterns, trajectories, and narratives— not univariate indicators and averages Develop rubrics that can include diverse evidence rather than large banks of tightly defined indicators	Show the expected contributions of different actors to the intended outcomes and impacts and to managing the risks of potential negative outcomes and impacts—not just those related to the specific intervention Have ways of zooming out and into elements, such as nested theories, so that details are included without being overwhelming Represent the interconn- ections between actors and between factors to identify 'levers and hubs'	Draw on multiple perspectives from different actors
Nonlinear relationships	Keep watch on patterns to identify when nonlinear change becomes important, such as feedback (positive or negative), thresholds, tipping points	Represent and communicate important nonlinear relationships	Continue to update the theory of change to identify thresholds, tipping points, and others

However, we will need to devise processes which allow for iteration without people feeling like they are not making progress. There will need to be a clear narrative about a theory of change not being *done* but a work-in-progress in ways that are useful rather than a source of ongoing stress and additional uncertainty.

In addition, the processes need to include drawing on complexity-appro- priate change theories. Instead of deterministic, behaviour modification,

stimulus-response models, which often implicitly or explicitly underpin theories of change, the change theories need to be more about facilitating change rather than producing it. A hydroponic farm, where the inputs and environment are strictly controlled, might lend itself to a traditional preplanned if-then theory of change; a diverse ecosystem with multiple species and microclimates interacting with each other and with changes in the seasons and the weather would need a change theory more about nurturing and stewardship.

For example, the underlying change theory of the Problem-Driven Iterative Adaptation to governance involves an iterative process involving four elements: (1) a local process of identifying priority problems to be solved; (2) supporting experimentation and positive deviance (identifying and learning from rare examples of success); (3) active, evidence-informed experimentation and adaptation; and (4) engaging champions to scale success (Samji et al., 2018).

The theory of transformation of Global Alliance for the Future of Food is described in terms of supporting transformative processes:

> Genuine food systems transformation takes place when diverse actions, networks, and individuals intersect across sector and issue silos, the global and local, the macro and the micro. These intersections facilitate convergence around shared visions and values and, ultimately, build critical mass and momentum behind tipping points that lead to healthy, equitable, renewable, resilient, inclusive, and culturally diverse food systems that dynamically endure over time. (Patton, 2020)

Conclusion

Table 10.1 summarises the implications of these features for using, developing, and representing theories of change. They will require many people to rethink in terms of how theories of change are used (as a nimble heuristic rather than a blueprint), how they are developed (through iterative, inclusive engagement), and how they are represented (in multiple complementary forms, including lists of principles, different views that zoom in and out, and metaphorical narratives).

These significant changes in our approaches to theories of change could enable more effective planning, management and implementation of initiatives in complex settings, as well as more useful and valid monitoring and evaluation. Given the high stakes of many of these initiatives—addressing crises including climate change, environmental collapse, and food security—these changes are urgently needed, as are the examples and the revised guidance on theories of change to support this.

References

CECAN. (2020). Complexity and what it means for policy design, implementation, and evaluation. Evaluation and Policy Practice Note No. 16. Retrieved from www.cecan

.ac.uk/wp-content/uploads/2020/08/EPPN-No-16-Complexity-and-what-it-means
-for-policy-design-implementation-and-evaluation-.pdf

Douthwaite, B., Kuby, T., van de Fliert, E., & Schulz, S. (2003). Impact pathway
evaluation: An approach for achieving and attributing impact in complex systems.
Agricultural Systems, *78*, 243–265.

Elliott, J. (2020). Evaluation with complexity in mind. [Doctoral dissertation, RMIT
University, Melbourne, Australia]. RMIT Research Repository. Retrieved from
https://researchrepository.rmit.edu.au/esploro/outputs/doctoral/Evaluation-with
-complexity-in-mind/9921904111801341

Gates, E. F., Walton, M., Vidueira, P., & McNall, M. (2021). Introducing systems—and
complexity-informed evaluation. In E. F. Gates, M. Walton, & P. Videuira (Eds.).
Special issue: Systems and Complexity-Informed Evaluation: Insights from Practice.
New Directions for Evaluation, 2021(170), 13–25.

Guerzovich, F., Aston, T., Levy, B., Chies Schommer, P., Haines, R., Cant, S., &
Faria Zimmer Santos, G. (2022). How do we shape and navigate pathways to social
accountability scale? Introducing a middle-level theory of change. CEDIL research
project paper 1. Centre of Excellence for Development Impact and Learning (CEDIL).
https://doi.org/10.51744/CRPP1

Ladner, D. (2015). Strategy testing: An innovative approach to monitoring highly
flexible aid programs. Working Politically in Practice Series, case study no. 3. The
Asia Foundation. Retrieved from https://asiafoundation.org/wp-content/uploads/2015
/10/Strategy-Testing-An-Innovative-Approach-to-Monitoring-Highly-Flexible-Aid
-Programs.pdf

Moore, G. F., Evans, R. E., Hawkins, J., Littlecott, H., & Melensez-Torres, G. J. (2019).
From complex social interventions to interventions in complex social systems: Future
directions and unresolved questions for intervention development and evaluation.
Evaluation, *25*(10), 23–45.

Patton, M. Q. (2020). The Global Alliance for the Future of Food adopts a theory
of transformation. Retrieved from https://futureoffood.org/insights/theory-of
-transformation/

Perrin, B. (1998). Effective use and misuse of performance measures. *American Journal
of Evaluation*, *19*(3), 367–379.

Samji, S., Andrews, M., Pritchett, L., & Woolcock, M. (2018). PDIA toolkit: A DIY
approach to solving complex problems. Retrieved from https://bsc.hks.harvard.edu/
tools/toolkit/

Part 4

Adapting theories of change for use in broader public policy contexts

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

11   Theories of change in evaluation of local government reforms

Kurt Houlberg and Olaf Rieper

Introduction and purpose

Evaluations of local government reforms have been sparse compared to how widespread such reforms are and their potential consequences for public services, politicians, and citizens. In this essay we use evaluations and research done on the Danish local government reform of 2007 to discuss the challenges evaluators face when evaluating such structural reforms and how theories of change have been applied, explicitly or implicitly.

The Danish local government reform in brief

The Danish government initiated the reform in 2002 and started a political process that involved the formation of a commission of civil servants and researchers in local governments (the Commission). The reform preparation was quite knowledge based, and the Commission referred to a multiplicity of studies of size, economy, and democracy. Some of the studies were commissioned by the Commission and some were available via the literature. Most were based on theories about economies of scale, including theories of scale advantages regarding professional capacity and specialization. Studies of size and local democracy were also included, although to a lesser degree.

The Commission delivered its report in 2004 with an overall conclusion that a reform was needed to improve professional sustainability, coherence in public service delivery, and utilization of economies of scale (Strukturkommissionen, 2004, pp. 24–28).

The reform reduced the number of municipalities from 271 to 98 and replaced the 14 counties with five regions. In addition, tasks were reshuffled across the three tiers of government, not least by transferring former county tasks to the municipalities (for example, specialized education for school children and specialized institutions for persons with disabilities). After the reform, the municipalities are responsible for the provision of most of the welfare services, except for the hospitals, which are the responsibility of the regions. The main purpose of the reform was to modernize the public sector and increase its effectiveness

by creating what were called *sustainable local governments*, to solve tasks of particular relevance to citizens and to secure synergy gains (Regeringen, 2004).

An official governmental evaluation of the reform was decided in 2011 after a period of political disagreement on the question. Also, a number of evaluations and research studies were launched independently of the government on the initiative of universities and research centres.

Applying theories of change in evaluations of local government reforms is a complex endeavour and raises huge challenges. In the following sections, we focus on three challenges facing evaluators of such reforms exemplified by the Danish reform and discuss how evaluators might begin to tackle the challenges.

Challenge 1: The political nature of the reform

Evaluations are inherently political and per se involve different stakeholders with different and often opposing views of both the justification, the meaning, and the results of the reforms (Pollitt & Bouckaert, 2011). Accordingly, decisions on evaluations of local government reforms include several potentially conflictual political decisions, such as whether a reform should be evaluated or not, and what should be evaluated, by whom, when, how, and for what purpose. The Danish reform was no exception and, only after several stages of political negotiations, a political agreement of an evaluation was finally reached in 2011 (Breidahl et al., 2017). Rather than a study of the extent to which the intentions behind the reform had been met or a study of its effects, the purpose of the evaluation was narrowed to a service check of four policy areas with a collaborative interface between municipalities, regions, and the state: (1) health, (2) specialized social tasks and special education, (3) nature and environment, and (4) regional development. The evaluation was organized as an internal evaluation—as are most evaluations of local government reforms (Swianiewicz et al., 2022)—and managed by an evaluation committee consisting of civil servants from the national government and representatives from the associations of municipalities and of counties. The responsibility for the evaluation thus lay with the stakeholders responsible for implementing the reform (Breidahl et al., 2017). The committee was not mandated to evaluate any democratic consequences of the reform. The evaluation of economic and professional sustainability was based on a survey asking mayors and CEOs whether they perceived the professional sustainability and administrative effectiveness of their own jurisdiction to be improved because of the reform (Ministry for Economic Affairs and the Interior, 2013, pp. 55–61).

Based on its analysis, the evaluation committee concluded that the municipalities and regions were in a process of realizing the objectives of the reform. The committee concluded that the reform strengthened the public service, both in relation to professionalism and economy; it also identified room for

improvement, especially where municipalities and regions collaborate about specialized social tasks (Ministry for Economic Affairs and the Interior, 2013, pp. 19–38).

The evaluation results were primarily used instrumentally to adjust the reform, but neither the decision to evaluate nor the evaluation process can be characterized by instrumental rationality (Breidahl et al., 2017). The evaluation process was more a negotiation process between stakeholders than an analytical process. Theoretically, it can be better understood based on interest and from an institutional cultural perspective designed to create meaning for the actors and to legitimize the reform (Breidahl et al., 2017). No theoretical reflections on the effects of local government size nor theories of change are found in the evaluation report.

In sum, the political nature of reform processes makes evaluation tasks difficult. The Danish case of an internal evaluation of the effects of the reform shows the limited space for a comprehensive theory-based evaluation.

There is no easy way for evaluators to tackle the political nature of the reform. The complexity of such reforms with years of negotiations makes the political investment in the reform huge and no politicians want to roll it back. A strategy might be to focus on certain elements of the reform instead of the reform as a whole. A similar strategy is recommended by Morkel (Essay 14) in the case of public sector planning in South Africa. Lima and Lafer (Essay 13) also point to the difficulties of using theories of change for a national multiyear development plan in Brazil.

Challenge 2: The complexity of the reform

The sheer complexity of such reforms given the multilevel characteristics of local reforms makes evaluation difficult. Across levels of government and organizations, many stakeholders need to be involved in the evaluation; there is a long path of causalities from the municipal political level through administration of services down to the single-service providers in schools, elder care centres, and others. The long path requires very complex theories of change. In addition, the municipalities are multipurpose organizations providing a variety of services and tasks (for example, social care, schools, labour market service, physical planning and infrastructure, environment, and water-supply). The theory of change will presumably be different for each category of services, and the evaluator must deal with several theories if the evaluation is to be comprehensive.

One way of handling this complexity is to limit the evaluation focus to parts of the reform or to certain kinds of effects. Examples from the Danish reform follow:

Besides the official evaluation, several researchers have evaluated various reform effects (for example, Bhatti & Hansen, 2019; Blom-Hansen et al., 2016, 2021; Lassen & Serritzlew, 2011; Hansen, 2013; Hansen & Hjelmar, 2015;

Kjær et al., 2010; Pedersen et al., 2022). These research studies are not compre-
hensive evaluations of the reform objectives but represent selected aspects of
the reform that the individual researchers have found theoretically interesting.
Methodologically, most of the studies are based on quasi-experimental designs
comparing the development in municipalities 'treated' with amalgamation with
a control group of municipalities not amalgamated. Theoretically, the studies of
economic aspects typically depart from an economy of scale logic, whereas the
studies of democratic aspects are nested in theories expecting negative demo-
cratic consequences associated with increasing jurisdiction size.

The studies find that economic robustness and fiscal management capacity have
increased in the amalgamated municipalities compared to the nonamalgamated
municipalities and that amalgamated municipalities have realized economies of
scale regarding the costs of administration and road maintenance. The cost savings
are offset by increasing costs for other services, and no economies of scale are wit-
nessed for total expenditures. Increased size is not reflected in higher effectiveness
and increased performance in core service areas, such as schooling. The amalga-
mations, in other words, do not seem to have improved professional sustainability.

Studies of democratic aspects show that the amalgamations have had nega-
tive consequences for the citizens on several dimensions, not least in the first
years after amalgamation. First, the citizens' internal political efficacy (that is,
trust in their own competences to understand local politics) and political trust
are reduced in the amalgamated municipalities. Apart from responsiveness, the
effects on internal political efficacy and political trust are less pronounced or
insignificant six years after the reform. Second, increased size reduces citizen
satisfaction with municipal services and with local democracy. Most of the iden-
tified negative democratic effects seem to be of a transient nature, and 14 years
after the reform, the effects have mostly vanished (Pedersen et al., 2022). The
studies also show that other dimensions of democracy, such as electoral turnout
and citizens' interest in politics, were not affected by the amalgamations.

Summing up, the positive effects for some of the economic aspects are
accompanied by losses for some of the democratic aspects. The uses of theo-
ries in academic studies are explicit, but they are theories of size and economy/
democracy and not theories of change addressing the change processes going on
in the implementation of the reform from the local levels to the operative levels
(schools and others).

The researchers have handled the complexity challenge by limiting the evalu-
ations to focus on selected parts of the reform and by mostly using quantitative
methods based on registers data and survey data.

Challenge 3: The long-time perspective

Because of the complexity of the local government reform and its structural
character, many consequences of reform will take a long time to be realized.

A long realization time requires theories that include changing contexts of the reform (for example, changes in national policy that influence the autonomy of the municipalities and changes in the population's local political interest and the population's expectations of the quality-of-service provision). It took more than 20 years to get the Danish local government reform from 1970 fully implemented, and several new laws and changes were introduced that affected the consequences of the reform during these decades. In addition to the causal issues of identifying the reform effects, the effects are also likely to be time dependent and therefore require conscious time reflections by the evaluator. The academic studies of the 2007 Danish reform, for instance, show that the short-run negative democratic effects seem to vanish over a longer time perspective.

Also, the ways evaluators might tackle the long-time perspective are similar to long-term evaluations in other fields. To handle the causality problem from changing contexts, contribution analysis, methods from the study of history (such as the study of documents, records and archives, and chronology), and the use of microdata are fruitful options. Because of the many interactions among elements of the reform and between the elements and the context, in-depth knowledge of the field must be included in the evaluation. Moreover, the conceptual use of evaluation results is probably more beneficial than in traditional short-term evaluations (Forss et al., 2021, p. 214).

Conclusion

Evaluations of structural reforms of local governments are embedded in political processes and are challenged by the political nature of local government reform, its complexity, and long-term implementation. The Danish case shows how evaluators have tackled such challenges by focusing on elements of the reform effects and by using macro data among other methods based on underlying theories of the effects of the size of municipalities. Although theories of the potential positive and negative effects of increased size are present, traces of theories of change are hard to find in the evaluations of the Danish reform.

In our review of research and evaluations, theories of change were seldom explicitly applied, but they were often implicit in the methodology of the studies. The academic studies of the effects of the reform focused on aspects of the reform and not on the reform as a whole. The main reason for this is probably the challenges that these studies encounter. The main challenges in our view are the sheer complexity of the reform: the municipalities as multilevel organizations with huge diversity of tasks, as well as the long-term perspective in implementing the reform down to the level of service provision. Moreover, the reforms require political investment, and politicians may recede from or avoid evaluations because a reform is cumbersome, if not impossible, to roll back. Learning might take place on the service level, not on the level of reform.

We see two complementary future avenues for enhancing the relevance and value of theories of change in evaluating complex policy reforms like local government reforms—a systemic and a partial theory of change route. A systemic route would imply elaborating more comprehensive and complex theories of change at the overall reform level to grasp and handle the multipurpose, multistakeholder, multifaceted, and multiactor nature of the reform. Such a 'grand' approach to the evaluation of the reform as a whole may, for instance, involve incorporating theories of systemic change (see Essay 12 by van der Knaap), an elaborated understanding of stakeholder interests and engagement (see Essay 15 by Olejniczak & Lyubashenko), and a more dialogical use of a theory of change (see Essay 4 by Dahler-Larsen). Given the complexity and limitations in using theories of change for the reform as a whole, *a partial theory of change route* for evaluating parts of the reform may be perceived as more realistic and beneficial in many reform evaluation scenarios. Such a partial route would imply designing and using theories of change for specific branches of services or for certain purposes and effects of the reform, including conscious reflections on the implications of focusing on selected reform elements, while leaving other aspects and effects behind a veil of darkness. Methodologically, either of these future avenues—the systemic theory of change route and the partial theory of change route—in our view would benefit from increased use of mixed methods and triangulation of a range of quantitative and qualitative data. This will help evaluators and policymakers to achieve a deeper understanding of the realization of reform objectives and the mechanisms by which the reform laid tracks (or not) for the reform as a whole or for selected aspects of the reform.

References

Bhatti, Y., & Hansen, K. M. (2019). Voter turnout and municipal amalgamations—Evidence from Denmark. *Local Government Studies*, *45*(5), 697–723.

Blom-Hansen, J., Houlberg, K., & Serritzlew, S. (2021). Jurisdiction size and local government effectiveness: Assessing the effects of municipal amalgamations on performance. *European Journal of Political Research*, *60*(1), 153–174.

Blom-Hansen, J., Houlberg, K., Serritzlew, S., & Treismann, D. (2016). Jurisdiction size and local government expenditure: Assessing the effect of municipal amalgamation. *American Political Science Review*, *110*(4), 812–831.

Breidahl, K. N., Gjelstrup, G., Hansen, H. F., & Hansen, B. H. (2017). Evaluation of large-scale public-sector reforms: A comparative analysis. *American Journal of Evaluation*, *38*(2), 226–245.

Forss, K., Lindquist, I., & McGillivray, M. (Eds.). (2021). *Long-term perspectives in evaluation: Increasing relevance and utility*. Routledge.

Hansen, S. W. (2013). Polity size and local political trust: A quasi-experiment using municipal mergers in Denmark. *Scandinavian Political Studies*, *36*(1), 43–66.

Hansen, S. W., & Hjelmar, U. (2015). Når kommuner bliver større: De korte og mere langsigtede konsekvenser for lokaldemokratiet. *Politica*, *47*(3), 103–123.

Kjær, U., Hjelmar, U., & Olsen, A. L. (2010). Municipal amalgamations and the democratic functioning of local councils: The case of the Danish 2007 structural reform. *Local Government Studies*, *36*(4), 569–585.

Lassen, D., & Serritzlew, S. (2011). Jurisdiction size and local democracy: Evidence on internal political efficacy from large-scale municipal reform. *American Political Science Review*, *105*(2), 238–258.

Økonomi- og Indenrigsministeriet. (2013). Evaluering af kommunalreformen. Marts 2013. København: Økonomi-og Indenrigsministeriet [Ministry for Economic Affairs and the Interior. (2013). Evaluation of the local government reform]. Copenhagen.

Pedersen, R. T., Kjær, U., Christensen, J., Hjelmar, U., Houlberg, K., & Petersen, N. B. G. (2022). *Lokaldemokratiet og borgerne—En analyse af borgernes syn på det kommunale demokrati 2001–2021*. VIVE.

Pollitt, C., & Bouckaert, G. (2011). *Public management reform* (3rd ed.). Oxford University Press.

Regeringen. (2004). *Det nye Danmark—en enkel offentlig sektor tæt på borgeren*. Regeringen, April.

Strukturkommissionen. (2004). Strukturkommissionens betænkning. bd.1: Hovedbetænkningen. Betænkning 1434. København: Strukturkommissionen. Retrieved from http://oim.dk/media/94820/Strukturkommisionens-betaenkning -hovedbetaenkning.pdf

Swianiewicz, P., Gendźwiłł, A., Houlberg, K., & Klausen, J. E. (2022). *Municipal territorial reforms of the 21st century in Europe*. Routledge.

12 Theories of change and the evaluation of sustainable impact

Moving beyond simplicity in development cooperation

Peter van der Knaap

Introduction

Theory-based evaluation offers an appealing approach for both evaluators and policymakers. This seems to hold true, especially in the field of development cooperation, where theories of change have long been an important part of the modus operandi. It is no surprise: policy theories provide focus for policymakers and allow for a logical starting point for important evaluation questions, such as: Did the intervention work? and Which improvements are possible?

Theory-based evaluation has also been the subject of fierce criticism, mainly because of its tendency to simplify societal challenges, public policy programmes, and the purpose of evaluation. With the adoption of the UN Sustainable Development Goals, this criticism is more relevant than ever: like other policy areas, development cooperation is shifting from individual projects to complex programmes and change on a national system level.

So, what does this rise of sustainability mean for the lure and relevance of theories of change and theory-based evaluation? This contribution starts from the still-strong promise of theory-based evaluation by revisiting its strengths and limitations (section 1). In section 2, the new millennium imperative that development policy must be sustainable is discussed. As policymakers are moving from simple projects to complex development interventions that aim to change entire social systems, evaluators are faced with growing complexity. Section 3 discusses alternative evaluation methods to evaluate effectiveness and makes the case for a new system-based theory of change approach. Herein, a theory of change not only serves to express the rather straightforward relationship between 'measure' and 'effect', but it also specifies the necessary system requirements for sustainable impact.

Section 4 concludes with a discussion of future directions of the evaluation of sustainable impact. Four avenues are explored: (1) the quest for sustainable development may be a new lifeline for the useful employment of theories-based evaluation by including system-level 'theories of sustainable change'; (2) there are useful, context-driven 'back tracing' and responsive methods available for system change or 'transition' evaluations; (3) the 'traditional' evaluation of

individual policy measures still has great value; and (4) when seeking sustainable development, policymakers should be beware of the risk of 'overreaching'.

1. Theory-based evaluation revisited

Decades after its inception, theory-based evaluation remains promising to both scholars and practitioners. In essence, it can be defined as the analysis and valuation of the contribution of a theory of change, 'intervention logic', or 'policy theory' to resolving or controlling society's problems. Herein, an intervention strategy is placed in rank and order to achieve goals—perfect for simple policy interventions.

Despite three decades of postmodernism, increasing policy complexity, and the rise of populism in Western democracies, policy and evaluation are still regarded as strong rational-analytic concepts. The rationale for public policy measures and, hence, public expenditure is to work systematically and even scientifically to achieve democratically legitimized goals. The usefulness of theories of change in evaluation and monitoring is present in all four phases of the policy cycle: (1) policy development, (2) debate and decision making, (3) implementation, and (4) rendering account. In all phases, policy-oriented learning and *focus* are important—and using theories of change provides just that (Van der Knaap, 2003).

To give an example, the theory of a simple programme may be: 'If we vaccinate children and provide clean drinking water and sanitation, then less children will die of disease before the age of five' (compare with USAid, 2022). Evaluation ex ante focuses on the expected causality and cost effectiveness of measures. In this way, using theories of change may enable better-informed, evidence-based decisions on policy proposals. Furthermore, a theory of change can be translated into an intervention-logic or implementation scheme, in which resources, actions, outputs, and results are combined. This allows for planning, monitoring, and management of implementation processes and progress but also for learning, including fundamental lessons about the validity of the theory itself.

Theory-based evaluation ex post revolves firstly around the assessment of the extent to which policy has resulted in success: Did the intervention work? Did it help to achieve the intended effect? Taking theories of change as a starting focal point for (summative) evaluation ex post creates a strong framework of reference and can boost the setup, implementation, delivery, and utilization of evaluation research. The second goal of theory-based evaluation is a better understanding of the underlying causal mechanisms. In such a case, the central question is: Were the assumptions on which the policy programme was based 'right' or not, and why? In this way, theories of change will serve both the notion of accountability and learning.

But like all theories, any theory of change is an abstraction of reality and will inevitably exclude a great wealth of facts and ideas—and deliberately so. In this

way, people get focus. The result, however, may very well be a simplified version of reality. Moreover, every such theory—and especially those that are well articulated—will direct perception, interpretation, and (thus) assessment.

It is difficult to underestimate the potential simplifying effect of a theory of change and the effects thereof on decision making, monitoring, and evaluation. Where too much emphasis is placed on anticipated causality and measurability, the result may be tunnel vision, rigidity, and even fear of and resistance to innovation (Van der Knaap, 2003). Everything that cannot be expressed in theories, performance data, and objectives may escape the attention of both policymakers and evaluators.

2. Sustainable development: From simple programmes to system-level complexity

In my 1988 student edition of *The Concise Oxford Dictionary* the word *sustainable* is not listed. *Sustain* is defined as 'hold up' and 'keep from failing or sinking, especially for a prolonged period', but the adjective itself is, rather surprisingly, lacking. Since the 1990s, however, *sustainable* has taken off in a spectacular way. In 2022, efforts to do good must lead to lasting results: road safety professionals to realize 'sustainable road safety' and innovation policy programmes to shift entire sectors into new products or services. This fundamentally changes the way policymakers and evaluators ought to consider 'effect', 'effectiveness', and 'impact'.

Let's take development cooperation as an example. Herein, the UN Sustainable Development Goals constitute 'bold and transformative steps… to shift the world onto a sustainable and resilient path'. It is no surprise that, in its new 2021 evaluation criteria, the Organization for Economic Cooperation and Development (OECD) included sustainability as a new evaluation criterion for development efforts. The main question perfectly summarizes the key point: Will the benefits of a policy intervention last? What is the extent to which the net benefits of the intervention continue, or are likely to continue? (OECD, 2021). According to the OECD, to assess sustainability, the evaluator must examine 'the financial, economic, social, environmental, and institutional capacities of *the systems needed to sustain net benefits over time*' (OECD, 2021, emphasis added).

In short, the rise of sustainability means the evaluation must consider resilient impact. This leads to 'tougher' questions, such as: Did we succeed in establishing a better 'enabling ecosystem' for the benefit we seek? or Will the benefit last when the policy programme is over?

The difference between implementing and evaluating a simple project (like a vaccination project) and a complex development intervention (like improving an entire country's health system for the future) is enormous. Simple projects work according to 'blueprint' approaches, producing standardized outputs with

relatively linear causal relationships between output and defined objectives (outcomes), within a specific timeframe (see Figure 12.1). In addition, agencies usually deliver a limited number of services to a defined and often small target population.

The Sustainable Development Goals' imperative that development and development measures must be sustainable—and therefore aim to change the underlying systems needed to eradicate poverty, secure human rights, and achieve climate objectives—means that policymakers and evaluators can no longer limit themselves to simple programmes but must aspire to (also) undertake complex development interventions.

As the OECD puts it: When the aim is sustainable change, the evaluation must take a broader perspective than individual measures and consider more fundamental, enabling changes in political, economic, financial, social, and environmental systems (OECD, 2021). In other words: evaluation must allow for complexity.

3. Alternative evaluation methods: Allowing for complexity

There is growing consensus that policymakers and evaluators alike must recognize the complexity of human behavior and the government and societal systems we create to live our lives (see Essay 11 by Houlberg and Rieper). There are vast numbers of individual elements at work that interact in complicated ways, whether it is ecosystems, energy networks, financial markets, or phenomena such as urbanization and migration (OECD, 2017).

The challenge of designing and applying the theory of change in such complex contexts is pronounced in many essays in this book including those by Morkel (Essay 14) and Lima & Lafer (Essay 13). Assessing outcomes of complex interventions poses special challenges for the evaluator (compare with USAid, 2022):

- Complex policy programmes may lack clearly defined activities and/or objectives and/or timelines.
- There are often multiple actors responsible for delivering activities and outputs.
- As interventions have a long-term horizon and a regional or countrywide reach, baselines and credible counterfactuals are often lacking.
- Even when evaluations include a counterfactual, the causes of observed changes may not be fully explained.

Figure 12.1 Model of a simple project

In evaluation literature, this has led to the development of alternative, more explorative and responsive methods that do not explicitly start off with a theory of change (Van der Knaap, 2011; compare with Essay 3 by Palenberg). Rather than the rational analytic 'golden standard' of randomized controlled trials (RCTs) or statistical (regression) analyses, the appraisal of causality uses a more constructivist approach. In addition, the importance of context is stressed. Examples of evaluation models that connect policy programmes to outcomes— or often vice versa—are outcome harvesting, modus operandi, realist evaluation, and contribution analysis:

- *Outcome harvesting* has specifically been developed to evaluate policy interventions in complex and dynamic contexts in which no simple theories of change can be constructed (compare with Wilson-Grau & Britt, 2013). The essence is that effects are not evaluated against objectives in policy programmes, but rather that the evaluator 'harvests' any change that has occurred in the domain of the programme. Once changes have been identified, an attempt will be made to trace the effects back to the measures undertaken under the programme. Lessons can be drawn on combinations of measures and conditions that 'work'.
- A similar approach can be found in the so-called *modus operandi* model (Scriven, 1974). Here, too, the objective is to—backwardly—infer causality between outcomes and programmes. Drawing from forensic science, the evaluator makes a list of possible causes for results and looks for evidence in a bottom-up fashion. Again, the aim is to draw conclusions on combinations or modes that produce results.
- The premise of *realist(ic) evaluation* is that evaluation is useful when it provides for answers to the question 'What works in which circumstances and for whom?', rather than merely 'Does it work'? (Pawson & Tilley, 1997). The evaluator seeks to identify the underlying mechanisms that explain 'how' outcomes were caused and the influence of context thereupon (compare with Douthwaite et al., 2017).
- Differently, *contribution analysis* starts with an articulation of the theory of change, but this may be complex rather than simple. Instead of RCTs, multiple lines of evidence are included to consider causal links in a concluding 'performance story' about what needs to be done in terms of conditions and interventions to achieve results.

The main advantage of these methods in 'the new age of sustainability' is that they make it possible to move beyond simplicity through the combination of methodological approaches that support the investigation and assessment of complex policy programmes in diverse circumstances. In all this, we must guard against pride. Policy, however well-intentioned, cannot do everything. In addition, it often remains useful to evaluate individual measures for

effects—as part of a comprehensive programme 'at system level' but also on its own.

4. Towards system-level 'theories of sustainable change'

As policymakers are moving from simple programmes to complex, system-level development interventions, it is time to ensure that evaluation does justice to increased complexity and system-level dynamics. Moreover, we must move towards a 'system theory-based evaluation' that embeds system approaches.

Herein, system change theories or *theories of sustainable change* are needed. Ideally, in the field of development cooperation, these higher level, more encompassing theories will make it possible to design policy programmes that will be 'self-supporting'. This means that all relevant beneficiaries buy into other coherent policy programmes and sustaining investments. The difference between simple intervention approaches and systems approaches is that this type of 'maintenance' is worked out and committed to in advance. In this way, results can be achieved for years to come.

This is important: as the OECD (2021) stresses, benefits often fade out after some time due to lack of maintenance and replenishment of new roads, bus fleets, teaching materials, hospital staff, and others. Systemic planning is necessary to have engineers trained, mechanics qualified, curriculum development experts and institutions in place, and so on. According to the OECD, 'The role of evaluation here can be to scrutinize assumptions in the theory of change for how sustainability is achieved' (2021, after Mansuri & Rao, 2012, p. 23; White et al., 2018, p. 24).

A theory of sustainable change not only serves to express the rather straightforward relationship between 'measure' and 'effect', but it also specifies the necessary underlying system requirements for sustainable impact. As in simple programmes, pathways can still be identified to connect these requirements to concrete conditions and actions. The difference is that there will be more than one relationship and that some will be conditional for others.

An example is the Expanded Program on Immunization, as described by Decouttere et al. (2021). Based upon a literature review, they constructed a 'conceptual diagram' in which the systemic objective to create 'an environment for health system development' is connected to the eventual impact of immunization though vaccination (see Figure 12.2).

A theory of change allows for focus but a system change theory or theory of sustainable change is undeniably more complex than a simple programme theory of change. Yet, without paying attention to underlying system requirements for sustainable impact, there can be no learning of lessons on the (likely) continuation of a policy intervention's benefits in the medium to longer term. These lessons may also highlight the potential scalability of the sustainability measures of the intervention within the current context or the potential replicability in other
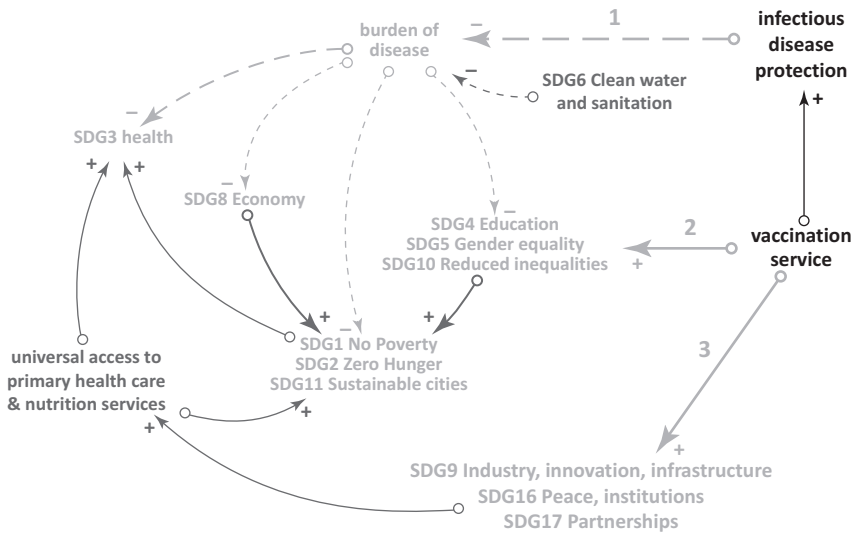
Figure 12.2 Example of a 'conceptual diagram' for the Expanded Program in
Immunization

contexts (OECD, 2021). They should also be used in making decisions about
discontinuation of programmes or aid: Is there an appropriate exit strategy that
can lead to a continuation of positive effects, or not?

As I see it, policymakers and evaluators that seek to make a sustainable
impact must adopt this broader perspective: efforts to improve health, safety, or
our climate should no longer be considered 'standalones'.

Both ex ante and ex post evaluation must examine the conditions for sustain-
able results. Ex ante, the question whether the necessary conditions for sustain-
able results can be met must be asked. Ex post, the first question is: Were those
conditions considered at the start of a policy programme and incorporated in the
design of the intervention and which lessons were learned along the way? The
second question logically is: Are the necessary conditions for sustainable results
in place and, hence, what can be said about the likelihood that investments in
development will yield the desired results in the medium to longer term?

Four future directions

Putting the notion of sustainability central in policymaking and evaluation is cru-
cial if we want to achieve and evaluate the effects, impact, and value-for-money of
policy programmes. After all, what is money spent worth when the net result will
be gone in just a few years after the intervention? For this reason, policymakers

and evaluators alike must take a system perspective, where a broader perspective is inherent to include the necessary conditions for sustainable results.

Despite the move from simple programmes to complexity, the promise of theory-based evaluation is still strong. True, a system perspective means more complexity, whereas one of the main advantages of theories of change and theory-based evaluation was to reduce complexity and provide focus. But even more sophisticated system change theories will enable policymakers to pay attention to the most important underlying causes and requirements for long-term successes. Even relatively complex theories of sustainable change can still provide a reference frame to stakeholders and other parties involved in the policy-oriented debate, allowing for well-considered argumentation.

These functions remain relevant to evaluators. Opting for the 'official' theory of sustainable change as the starting point for research and assessment may not only offer a logical starting point for assessing success, but it may also offer a good opportunity to connect with policymakers' thoughts and ambitions (whether publicly expressed or not).

The rise of sustainability means that both policymakers and evaluators must face up to more complex causal relationships between policy and benefits and, hence, use the alternative methods described to establish an evidence base between complex sets of interventions and outcomes. In addition, they must establish a 'higher level' orientation towards relevance and learning.

My conclusion is: Theories of change and theory-based evaluation are here to stay. They can be improved by including system change theories or theories of sustainable change. I see four future avenues for a world in which sustainable impact is what is wanted:

1. When we succeed in working towards theories of sustainable change and 'system change theory-based evaluation', the quest for sustainable development may in fact be a new lifeline for a useful employment of theory-based evaluation.

   In complex fields like development cooperation, road safety, or environmental protection, policymakers increasingly opt for a 'systemic approach' as an intervention strategy. Herein, various policy measures and actors must 'work together', while taking account of contextual developments. These 'system approaches' must be made part of theories of change and policy evaluation. At the same time, people still need focus to organize their thinking, knowledge management, debates, and decision making. With system change theories we do justice to complex policy programmes and acknowledge the contributions from various network partners.

2. There are useful context-driven, 'back tracing' methods available for evaluating system change or 'transition'.

Policies that seek to achieve sustainable change—whether in development cooperation, innovation policy, or road safety—require a system-based approach that is, by definition, more complex than the implementation of individual measures. As policymakers are moving from 'simple projects' to 'complex development interventions', evaluators are faced with growing complexity. They must be able to employ these alternative methods to do justice to contextual factors that support (or undermine) the sustainability of benefits, whether these can be assessed quantitatively or qualitatively. These include but are not limited to stakeholder ownership and engagement, absorptive capacity, political will, and long-term resource availability.

3. Within complex policy interventions, the deployment of 'traditional', RCT-like evaluations of individual policy measures still has great value.

The evaluation of individual policy measures is still useful, but mainly as part of 'bigger', sustainability-oriented evaluations on a system level. In *Less Pretension, More Realism* (IOB, 2019), the main finding of the evaluation was that two important programmes achieved several tangible outputs; but in the longer term, their impact was limited. While concrete results were achieved (such as the establishment of peace committees or infrastructural works), outcomes were often restricted to individual cases or were local in scope. In only a few instances did outcomes 'trickle up' and positively influence broader socioeconomic development or reduced levels of conflict or insecurity.

4. When seeking sustainable development, policymakers should be aware of the risk of 'overreaching'.

The ever-increasing complexity of social and environmental problems is reflected in policy programmes. As the IOB evaluation demonstrates, it is useful to base interventions on explicit and tested assumptions and a sincere appraisal of the sustainability of expected outcomes. The conclusion may well be that more realism is needed: be aware of the risk of 'overreaching'. Given the many variables that determine long-term success on a system level, policymakers should beware of 'optimism bias'.

References

Decouttere, C., De Boeck, K., & Vandaele, N. (2021). Advancing sustainable development goals through immunization: A literature review. *Global Health*, *17*(1), 95.

Douthwaite, B., Mayne, J., & McDougall, C. (2017). Evaluating complex interventions: A theory driven realist-informed approach. *Evaluation*, *23*(3), 294–311.

IOB. (2019). *Less pretension, more realism*. Ministry of Foreign Affairs of the Netherlands, Policy and Operations Evaluation Department, The Hague.

Mansuri, G., & Rao, V. (2012). *Localizing development—Does participation work?* The World Bank.

OECD. (2017). *Debate the issues: Complexity and policy making*. OECD Insights, OECD Publishing. http://dx.doi.org/10.1787/9789264271531-en

OECD. (2021). *Applying evaluation criteria thoughtfully*. OECD Publishing. https://doi.org/10.1787/543e84ed-en

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage Publications.

Scriven, M. (1974). Maximizing the power of causal investigations: The modus operandi method. In W. J. Popham (Ed.), *Evaluation in education: Current applications* (pp. 68–84). McCutchan Publishing.

USAid. (2022). *Evaluating complex programs: Complex interventions and the challenges of evaluating them*. Module 17. USAid Learning Lab.

Van der Knaap, P. (2003). Theory-based evaluation and learning: Possibilities and challenges. *Evaluation*, *10*(1), 16–34.

Van der Knaap, P. (2011). Sense and complexity: Initiatives in responsive performance audits. *Evaluation*, *17*(4), 351–363.

White, H., Menon, R., & Waddington, H. (2018). *Community-driven development: Does it build social cohesion or infrastructure? A mixed-method evidence synthesis*. Working Paper No. 30. International Initiative for Impact Evaluation.

Wilson-Grau, R., & Britt, H. (2013). *Outcome harvesting*. Ford Foundation Mena Office.

13 Use of theory of change as a management tool for government multiyear development plans

The case of Brazil's Federal Development Plan

Lycia Lima and Marina Lafer

Introduction

Theory of change is a classic programme theory evaluation tool that emerged at project-level interventions to understand 'how change happens' (Vogel, 2012), by linking inputs and actions to plan, track progress, and assess how these lead to final outcomes. Several governments, such as Brazil, Canada, Chile, Cape Verde, Colombia, and Mexico, as well as countries in the African continent pointed out by Morkel (Essay 14), such as Benin, Ghana, Lesotho, Namibia, South Africa, Uganda, and Zambia, adopted theory of change, as well as other programme theory tools, to plan, monitor, and evaluate interventions.[1]

With the adoption of the UN Sustainable Development Goals, there has been a global movement of shifting the use of theory-based tools from individual projects to complex programmes and national system levels (Van der Knaap, Essay 12). Brazil is an example of a federal government that recommends the use of programme theory tools to design multiyear development strategies (that is, beyond their usual intervention-level scope). It is unclear, however, whether its application is useful to plan, track, and assess change in these contexts, due to high levels of complexity as discussed below.

Using Brazil's 2020–2023 national multiyear development plan, we analyse whether theory of change is sufficient, feasible, or adaptable enough to suit this

---

1 The term *intervention* will be used to refer in a general manner to projects, programmes, and policies, despite these being different concepts. Projects are single interventions implemented in one or more locations; programmes are several related activities or projects implemented to contribute to a common goal; and policies are sets of norms, guidelines, or rules that guide decisions of organisations in a certain area, which may be materialised through several programmes (Morra Imas & Rist, 2009). On the other hand, the concept of programme is used differently in the context of Brazil's multiyear development plan (abbreviated as PPA). PPA-level programmes differ from interventions, as will be seen in this essay; the former are macro-thematic areas that potentially, but not declaredly, nest interventions within the Brazilian public sector planning process. PPA-level programmes guide the definition of broad targets as well as resource allocation in the medium term.

purpose while exposing the difficulties of using it to unveil the 'black box' in aggregated, multi-intervention settings.

Theory of change use in government multiyear development plans

Theory of change is both a process and a product widely used in international development, public administration, and civil society. It provides insights to support design, strategy, implementation, evaluation, communication, and learning around interventions that seek to alleviate, mitigate, or solve development problems (Vogel, 2012). Generally, it presents a visual systematisation of the logical chain of events to describe change (Vogel, 2012) and is represented by steps such as inputs, activities, outputs, outcomes, and impacts. Theory of change takes into consideration the contextual conditions in which a problem manifests itself and the assumptions to fulfil the cause-effect relationships between its elements (Morra Imas & Rist, 2009).

Starting as a project-level framework as part of the programme theory area of evaluation thought, theory of change 'promises' to help programme managers reach the intended long-term changes. For this reason, governments might find it useful when developing national long-term strategies, even though 'the difference between implementing and evaluating a simple project (like a vaccination project) and a complex development intervention (like improving an entire country's health system for the future) is enormous' (see Essay 12 by Van der Knaap).

Brazil is one of the national governments that has its own multiyear development plan, named Plano Plurianual (PPA), that is used to establish strategies and priorities and coordinate actions across sectors to reach development objectives (United Nations, 2018). Issued as a law on a quadrennial basis, PPAs generally define principles, objectives, goals, and budgets by 'PPA-level programmes', which consist of macrothematic areas that guide the definition of broad targets as well as resource allocation. For instance, the education sector in the latest federal PPA (2020–2023) has six 'PPA-level programmes' (Early Childhood Education, Quality Basic Education, Professional and Technological Education, Higher Education, Special Needs Education, and Educational Statistics and Assessments) with respective objectives and goals (Brazil, 2019a, Annex I).

Brazil's federal government's latest supporting materials to guide sectors during the PPA 2020–2023 planning cycle advocate for the use of programme theory tools, such as the logical framework approach (logframe) and theory of change for each PPA-level programme in its final form (Brazil, 2019b). However, even though the use of programme theory tools is advocated to assist in planning, monitoring, and evaluating Brazil's PPA, we were unable to find any practical applications of these; their application and usefulness are, thus, not evident. We will illustrate an exercise for a slice of the national education development strategy as

an example to analyse the suitability of building a theory of change in complex contexts.

Analysis of the suitability of the use of theory of change for PPA 2020–2023

Both PPA law (Brazil, 2019a) and the accompanying technical manual (Brazil, 2019b) provide information and guidance to match PPA elements to programme theory components. According to these documents, the PPA category level 'Principle' (*Diretrizes*, in Portuguese) is analogous to 'Impact' in the theory of change framework. For the basic education example, its 'Principle' in the latest PPA is 'priority dedication to quality of basic education, especially early child-hood education, and preparation for the labour market'. Likewise, PPA categories 'Programmes and Their Objectives', 'Indicators', and 'Goals' (*Programas com seus Objetivos*, *Indicadores*, and *Metas*, respectively) are equivalent to 'Outcomes'. For basic education, the intended outcome for the current period is to 'raise the quality of basic education, promoting access, permanence and learning with equity' and its target indicator is to 'achieve the 5.59 goal in the Synthetic Basic Education Index (Ideb)'. Moreover, 'Budgetary Actions', 'Non-Budgetary Actions' and 'Multi-Annual Investments' (*Ações Orçamentárias*, *Ações Não-Orçamentárias*, and *Investimentos Plurianuais*) correspond to 'Outputs, Processes, and Inputs' in the theory of change framework (Brazil, 2019b, p. 27).

Nonetheless, the output-level entries listed in the PPA describe actual inter-ventions (policies and programmes) that could themselves be characterised as complex processes with their own specific theories of change; due to the multiple layers of theories of change for complex interventions, there is a great risk of simplifying reality and leaving important information out for both policymakers and evaluators (Van der Knaap, Essay 12). For instance, the output 'National School Feeding Programme' certainly undertakes activities such as fiduciary transfers to subnational levels, development of nutritional norms, training of nutritionists, and more; these are not explicit in the PPA and, consequently, in the macro-level theory of change. One potential explanation for not listing all the activities in place for each goal may be that, as PPA is a planning tool, it is published in advance, at a time that specific sectoral policies have not yet been determined. Another potential explanation is that spelling out all details of the programmes listed as activities and outputs to achieve the ultimate development goal may lead to an enormous amount of information that may turn the theory of change into an unworkable tool.

As a management tool, a given theory of change is the basis for building a monitoring plan that allows for assessing whether the intervention is on track with reference to planned activities and associated outputs and outcomes. The policies and programmes listed under basic education, when included in the

2020–2023 PPA theory of change at the output level, end up being 'black boxes' themselves. All these activities should be detailed in a theory of change for it to serve as a management tool, as these need to be monitored to ensure that the basic education PPA programme is being executed in an adequate manner.

If the outcome for Quality Basic Education is not met—that is, the Synthetic Basic Education Development Index is below the target—it would be impossible to identify what is not working based on the macro-level PPA theory of change. For instance, it could be that the 'Granting of Scholarships to Support Basic Education' was poorly executed due to disbursement delays within a certain region; it could also be that the same intervention was perfectly implemented, but the way it was designed simply does not lead to the intended change (for example, the amount granted might be insufficient). Without being able to establish intervention-level indicators on their inputs, activities, and outputs, for instance, it is not possible to conduct a sense-making analysis on the components and links of the theory of change; there is no explicit evidence to identify which components are favouring or hindering the achievement of the outcome target.

Even though it may be somewhat possible to identify what the government plans to achieve in terms of education (Principle), based on the macro-level PPA theory of change, it remains unclear how it will get there. Theory of change can be used as a planning tool at the PPA-level; however, without knowing what is 'inside' each PPA-level programme, it cannot be used as a management tool. It would require particularly robust assumptions to be able to conclude that the 'inner workings' (Connell & Kubisch, 1998, as cited in Freer & Lemire, 2019), that is, the interventions and their respective theories of change, are working effectively.

Limitations of the theory of change for government multiyear development plans

A broad, generic theory of change organised by 'umbrella' themes (PPA-level programmes) towards a broader 'goal' can be used to communicate the bigger picture across stakeholders and create consensus around priorities and long-term vision. Nonetheless, the existence of supporting orientation materials to apply the theory of change tool for the management of multiyear development plans is insufficient and remains unclear in practical terms.

To understand whether and how change is happening, it is important that theories of change bring additional information. This includes which goods and services are expected to be offered to people or organisations, what resources and activities are needed, which departments are responsible to put the interventions in place at the country's national and subnational levels, and whether the interventions will require local adaptation, and so on.

The federal government's guideline on programme theory tools (Brazil, 2018) might be useful for planning, resource allocation, and targeting ground-level

interventions. Nonetheless, it might not be suitable when using programme terminology at the PPA level given its distance from day-to-day implementation and might represent a deliberate 'simplification of what is clearly a far more complex system' (see Essay 14 by Morkel).

Steps towards moving forward

To deal with the limitations that generic theories of change bring as a planning, monitoring, and evaluating tool, several theories of change for local-level, project-based interventions under these generic depictions might be an alternative to be investigated. For that, overarching theories of change for government multiyear development plans could be drilled down into 'nested theories of change' (Mayne, 2015). This would help to unpack complexity and make it more likely to comprehensively think through and identify the most appropriate activities, outputs, and relevant stakeholders at local levels while integrating these into a broader picture. The means to achieving macro-level changes are less likely to fall into a generic, vague theory of change that probably would not indicate what needs to be done and to happen.

To fill in the macro-level, overarching theory of change 'black box', there must be a definition of what interventions should be in place at the local level. Thinking education-wise, several national and subnational agencies are responsible for planning and executing interventions related to the themes listed in the PPA-level programmes. These intervention-level theories of change should communicate with each other and, together, feed into a 'parenting' theory of change to achieve education's long-term goal.

Several challenges might be encountered when putting these steps into practice moving forward; these are related to the design of the theory of change, to the resources required to conduct it, and to the processes used to facilitate this exercise. Even though creating nested theories of change might seem like a possible solution, it remains questionable whether creating overarching and respective nested theories of change allows for intervention planning and visualisation, tracking, and evaluation at such macro levels of interventions.

Stakeholder engagement is also key for the success of a theory of change, since it increases the likelihood that context, programme knowledge, and diverse perspectives are considered (see Essay 15 by Olejniczak & Lyubashenko) and key responsibilities are assigned for change to occur (Lam, 2020). Moreover, when the theory of change object is a four-year, national strategy that involves multiple themes, layers of interventions, and actors at multiple government levels, it seems difficult from articulation and operational standpoints to involve all stakeholders in developing a cohesive nest of theories of change. Capacity is another challenge for the suitability of this process (Lam, 2020; Mason & Barnes, 2007; Vogel, 2012); stakeholders widespread at local levels would need to be familiar with the approach to avoid vague and narrow theories of change being developed.

Last, intervention theories of change should fit under an overarching PPA-level programme theory of change in an integrated, sense-making, and complementary manner so that they do not undermine the success of one another. It will be challenging for central levels of government to put together all the nested theories to make sense in reaching the intended change sought in the broader picture.

Closing remarks

Theory of change is a tool that has been used to understand whether and how change happens in development interventions. While its suitability to serve as a tool to plan, monitor, and evaluate change in macro, multiyear development strategies has been formally advocated by governments, its application has not yet been observed.

Using the basic education PPA in Brazil as an example, we pointed out that—in an attempt to unveil a 'black box' regarding the actions the Brazilian government is taking to achieve the intended changes—several other 'black boxes' surfaced. Even though the federal Brazilian government states the goals it wants to achieve, the roadmap that needs to be followed is still uncertain and difficult to put together.

Nested theories of change, when conducted through coordinated and participatory processes, can potentially help to fill in these 'black boxes' for complex contexts such as those of multiyear development strategies. However, when thinking of articulation and capacity at national, regional, and local levels, especially in a continental-size country such as Brazil, the suitability of macro theories of change as a management tool for this purpose remains challenging and must be assessed on a case-to-case basis. If articulation and capacity challenges to put together a nested theory of change for this purpose are greater than its benefits, these might easily fall into simplifications of reality and thus not be useful management tools. In this case, intervention-based theories of change may be more effective for management purposes, even though these may not be able to depict the full, complex picture.

References

Brazil. (2018). *Avaliação de políticas públicas: guia prático de análise ex ante* [*Evaluation of public policies: Practical guide on ex ante analysis*]. Volume 1. Casa Civil & IPEA.

Brazil. (2019a). Lei n° 13.971 de 27 de dezembro de 2019: *Institui o Plano Plurianual da União para o período de 2020 a 2023* [Law no. 13.971 of 27 December 2019: *Establishes the National Multiannual Plan for the 2020–2023 period*]. Retrieved from www.gov.br/economia/pt-br/assuntos/planejamento-e-orcamento/plano-plurianual-ppa

Brazil. (2019b). Manual técnico do plano plurianual do governo federal 2020–2023 [Technical manual of the 2020–2023 federal government multiannual plan]. Ministério

da Economia. Retrieved from www.gov.br/economia/pt-br/assuntos/planejamento-e -orcamento/plano-plurianual-ppa/arquivos/manual-tecnico-do-ppa-2020-2023.pdf

Connell, J. P., & Kubisch, A. C. (1998). Applying a theory of change approach. In K. Fulbright Anderson, A. C. Kubisch, & J. P. Connell (Eds.), New approaches to evaluating community initiatives (Volume 2): Theory, measurement, and analysis (pp. 15–45). Washington, DC: The Aspen Institute. https://www.semanticscholar.org/ paper/Applying-a-Theory-of-Change-Approach-to-the-of-and-Kubisch/b5ae8eceb15 ffaf10e2268af2a045c62b4e13f51#citing-papers)

Freer, G., & Lemire, S. (2019). Can't see the wood for the logframe: Integrating logframes and theories of change in development evaluation. *Canadian Journal of Program Evaluation/La Revue canadienne d'évaluation de programme* 33(3), Special issue/ Numéro special, 336–353. https://doi.org/10.3138/cjpe.53007

Lam, S. (2020). Toward learning from change pathways: Reviewing theory of change and its discontents. *Canadian Journal of Program Evaluation/La Revue canadienne d'évaluation de programme*, 35(2), 188–203. https://doi.org/10.3138/cjpe.69535

Mason, P., & Barnes, M. (2007). Constructing theories of change: Methods and sources. *Evaluation*, *13*(2), 151–170. https://doi.org/10.1177/1356389007075221

Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, *30*(2), 119–142. http://dx.doi.org/10.3138/cjpe.230

Morra Imas, L. G., & Rist, R. C. (2009). The road to results: Designing and conducting effective development evaluations. World Bank. https://openknowledge.worldbank .org/handle/10986/2699

United Nations. (2018). *The 2030 agenda and the sustainable development goals: An opportunity for Latin America and the Caribbean*. LC/G.2681-P/Rev.3. ECLAC. Retrieved from www.cepal.org/en/publications/40156-2030-agenda-and-sustainable -development-goals-opportunity-latin-america-and

Vogel, I. (2012). *Review of the use of 'theory of change' in international development*. UK Department for International Development. Retrieved from www.theoryofchange .org/pdf/DFID_ToC_Review_VogelV7.pdf

14 Theories of change in complex macro public-sector planning settings in Africa

How useful are they?

Candice Morkel

Introduction

Over the past four to five decades, programme theories have been widely used in international development practice, where the value of logic modelling and theories of change has been widely recognised (Funnell & Rogers, 2011; Rogers & Weiss, 2007, p. 64). For evaluators, a theory of change is the *sine qua non* of any evaluation, programme, policy intention, or development intervention. The 'theory of change' has become part of the *lingua franca* of evaluators and nonevaluators alike, and it is common to see these integrated into the planning and measurement frameworks of civil society organisations, development programmes, as well as social enterprises (the latter in the wake of the rapid growth of impact investment, management, and measurement). The conflation of various terms associated with theories of change such as *results framework*, *logframe*, *results chain*, and so forth has also resulted in some confusion and inconsistencies in the usage of the concept.

Recently, the prominence of theories of change in better understanding complex interventions and how to attribute causality has also been on the rise (McConnell, 2019, p. 214; Murphy & Jones, 2021). There has also been a deepening recognition of the importance of theories of change in the planning and monitoring and evaluation (M&E) systems across governments on the African continent (including Ghana, South Africa, Benin, Uganda, Zambia, Lesotho, and Namibia) and an upsurge in their formal adoption as part of public-sector planning requirements. For example, national compliance frameworks in South Africa require the inclusion of theories of change to guide planning, and planning templates require the articulation of outcomes, outputs, indicators, and targets (SA DPME, 2021).

It is not surprising that African governments have incorporated the use of theory of change to frame their planning, as globalisation has exposed all states and bureaucracies to the form, functions, and mechanisms of modern democracies, and has facilitated their assimilation into the normative rules of governance. Public-sector transformation efforts; the pursuit of good governance; the politico-economic principles associated with modernisation; and the demands for efficient, capable, and accountable bureaucracies are the hallmarks of the

wave of new public management that swept across Europe and found its way to African bureaucracies.

There is a growing realisation of the importance of context in the international development sector in particular, and the wholesale adoption of a one-size-fits-all approach around policy, planning, implementation, and governance in pursuit of development outcomes is steadily waning. Therefore, a critical assessment of the value and applicability of various technical approaches to planning, prioritisation, decision making, and governance in the African public-sector context in general is needed. Moreover, these should go as far as assessing their applicability to individual African states' political, cultural, and socioeconomic contexts.

The specific history and rationale of the adoption of planning, prioritisation, and decision-making models in African democratic states is related to the institutionalisation of Western forms of bureaucratic functioning, which was widely exported to post-independence democratic African states. Certain technical governance support mechanisms—such as results-based planning and management, performance management, measurement, accountability, and transparency—have been widely accepted as ideal and desirable. They are seen as indicators of a stable democracy and are signals of progress and development to the global community, accruing such dividends as improved investment ratings, greater access to foreign investment, and more attractive aid conditions. However, in the context of an increasingly complex world, with multiple, intersecting crises and wicked problems, governments need to be aware of new developments around systems-oriented theories of change and be more critical of linear logic models that had dominated earlier conceptualisations of planning and measurement of development outcomes. Such earlier approaches and tools continue in popularity, despite what we now know about their weaknesses in adequately reflecting complexity and without sufficient recognition of their limitations in determining development results.

The limitations of the use and application of theories of change in the public sector

The boundaries and limitations of the use and application of theories of change in the public-sector context (particularly in complex systems, with multiple intersecting interventions in a large organism such as the state) have been under-examined. There is insufficient examination of the extent to which a theory of change can map out a predictable pathway of success in macro development plans of the state in Africa, vis-à-vis how they work in programme contexts. In many public-sector planning scenarios, the distinction between the design elements and processes that work for programme theory—and the process of de/reconstructing, depicting, and developing a pathway of change for complex systems—has not been made.

To a certain extent, it appears that the growth of evaluation capacity-building programmes, particularly in Africa, has contributed to the increasing prominence

of theories of change and a greater recognition of their value in planning, monitoring, and evaluation (Bakken et al., 2015; Basheka, 2016). In efforts to address the capacity constraints in the public sector in evidence use for decision making, building the capacity of public officials has become a key feature of international development programmes, and global initiatives have been established to specifically address the issue of capacity development in national evaluation systems (see for example the National Evaluation Capacity initiative https://nec .undp.org/about-nec). Training programmes—particularly foundational training in M&E—include, as a staple, training on the theory of change. The increasing integration of the theory of change in public-sector planning, even amongst those who might have initially shunned the language of monitoring and evaluation as alienating or an artefact of the M&E industry, means that some of the resistance has fallen away. A theory of change has increasingly been accepted as the fundamental starting point of all planning processes and a litmus test of goal achievement in the public sector in countries such as South Africa (SA DPME 2019, 2021).

One of the challenges is that there might be an oversimplification of the way in which the knowledge and skills required for building theories of change are developed in these contexts. There are no standards for curriculum design and content for M&E training programmes, and therefore the time spent on building the theory of change could span anything from hours to days or weeks. In the era of the COVID-19 pandemic, many M&E training courses had to be truncated, resulting in an even shorter period spent on topics such as theory of change. What is mostly introduced to participants in such training programmes is the construction of a simple linear, logical pathway of change along the lines of inputs, activities, outputs, outcomes, and goals or impacts.

The logical simplicity of the boxes and arrows that are usually used to demonstrate the pathway of change towards impact may be appealing to public officials who are faced with the complexity of the public policy and development mandate that they face. Depicting how public-sector interventions will deliver major impact in areas such as job creation, unemployment, climate change mitigation, and poverty can be very attractive. Public-sector leaders are usually expected to demonstrate the impact of their work across short administrative terms, generally around five years, and the depiction of a linear progression of boxes and arrows from activities to results is promising. However, often little attention is paid to the spaces between the boxes or shapes—the mechanisms of change—and in many cases even the risks and assumptions are omitted. The intersection of a single results chain with other state- and nonstate-led interventions and results chains, as well as the systemic interactions between them and the prevailing context, is often not part of the picture. Furthermore, the theory of change may have no actual relationship to the government's implementation plans, as in many cases there is no alignment between what has been stated as an intention in the theory of change and what is contained in budget and

implementation plans. Rogers and Weiss (2007, p. 64) identified this problem more than 15 years ago, noting that many organisations have claimed to adopt programme theory, but in fact they have introduced only a semblance of it in the form of an implementation plan, depicting through the boxes and arrows how inputs portend to transform into impacts.

Rationale for a differentiated approach

There are a number of reasons why a more nuanced approach to the theory of change needs to be considered in the public sector. In the first instance, consideration must be given to the possibility that a form of isomorphism is at work in the growing demand for the construction of national M&E systems. As alluded to previously, the ubiquity of modernised democratic systems of governance has the potential for national M&E systems to take on a teleological form. This can be circumvented by a critical, comprehensive, and in-depth M&E systems analysis, led by local partners, to examine what is needed ahead of the adoption of any tools or solutions and to adapt these to the local context and needs. For example, logic modelling that might work in simple project-based interventions might need to be significantly adapted to account for complexity and the demands of public-sector policy and governance. A compliance- and performance-oriented environment has often led governments to pursue simpler goals and outputs, rather than transformational development goals that seek to change the livelihoods and well-being of citizens. This has created a perverse incentive to select 'low-hanging fruit', or to downgrade targets to achievable and monitorable outputs and deliverables for which an auditable portfolio of evidence may be established, as this is what is valued in a performance-driven, accountability culture.

Secondly, the political context of government decision making needs closer examination. Insufficient attention has been paid to gaining a deep understanding of existing policy and budget decision-making systems and processes before adopting the technical aspects of M&E systems, including the use of theories of change to drive planning and evaluations. As pointed out almost two decades ago by Kusek and Rist (2004), M&E 'is a political process with technical dimensions'. When introducing technical tools, models, and approaches into these systems, they are often superimposed onto existing systems and become tangential, acting as parallel systems to the *realpolitik* of decision making and prioritisation. Although it is well documented that the political and situational context must be taken into consideration when establishing government M&E systems, conceding to this is often an act of instrumentalism rather than a genuine consideration of how the evidence production and decision-making machinery of the institution needs to adapt to (or be adapted to). Limited timeframes, such as those determined by technical partnership contracts and agreements or the demands of tight policy and budget cycles, pose a constraint to undertaking

the broad consultative and reflective processes required to build contextually tailored solutions. The available technical solution or model (for example, the traditionally linear model of a theory of change) is therefore incorporated into the system without the necessary in-depth engagement, analysis, and adaptation.

The challenge of poor intergovernmental relations as well as weaknesses in integrated governance are also underexamined in the introduction of technical fixes (as M&E systems are accused of being) into the state. It might be argued that theories of change, as programme theory, could only conceivably work for simple government interventions or programmes, given the complexity of public good delivery and the many intersecting policy intentions, development goals, delivery mechanisms, agents, intervening variables, stakeholders, risks, and assumptions. A theory of change encompassing macro goals, such as improving health or education, could become unwieldly in order to accommodate the mechanisms of change for a multiplicity of programmes and projects, as well as the various aspects listed previously. It is not clear whether theories of change are able to viably depict causal chains for large, complex systems without sacrificing many of the finer details around change mechanisms.

The South African National Ministry of Basic Education, for example, states in its Strategic Plan 2020–2024 that 'a theory of change for the entire basic education system will inevitably be a simplification of reality' (SA DBE, 2020, p. 24) and states that the figure depicting the theory of change 'is deliberately a simplification of what is clearly a far more complex system'.

Theories of change and the South African planning landscape

The introduction of results-based M&E to the planning landscape of the South African public administration provides evidence of how theories of change have been institutionalised in public-sector planning. However, it is not the purpose of this essay to provide a comprehensive history of the country's planning and M&E landscape, as others have addressed this in greater detail in a few key studies and scholarly works (see, for example, Cloete et al., 2014).

To begin with, the South African government's macro planning processes are informed by extensive party-political and stakeholder consultations. The national Department of Planning, Monitoring and Evaluation (DPME), supported by the national treasury, is the custodian of the technical planning processes where these policy pronouncements find expression in government service delivery plans, guided by various regulatory frameworks. These regulations are underpinned by key legislation, such as the Public Finance Management Act, Act no. 1 of 1999 as well as the Public Audit Act, Act no. 25 of 2004.

These frameworks are strongly influenced by principles of results-based management and employ the framing that has been made popular by texts such as *Ten Steps to a Results-Based M&E System* by Kusek and Rist (2004). The latest planning guideline (SA DPME, 2019) proffers that 'Theory of Change is

one of the tools that can be used to determine a pathway for achieving desired results. Elements of the theory include the pathway of change; results and pre-conditions (impact, outcomes, outputs, activities, and inputs); indicators; and assumptions'. Whereas before there was little recognition of the importance of designing pathways of change in government programming, it is an achievement that programme managers, and indeed other public officials in the bureaucratic system, are increasingly familiar with the terminology associated with M&E systems, including theory of change. Even though all may not agree with it, there is widespread compliance with the previously eschewed M&E tools and techniques as part of the planning landscape.

Overcoming the challenges: The viable use of theories of change in public-sector planning and M&E

In order to ensure that theories of change do not remain a techno-fix, and evidence-informed decision making including the capacity to plan, prioritise, produce, and use evidence to guide policymaking is deepened, some challenges need to be overcome. For example, although official guidelines in the South African example propose various methods such as problem tree and fishbone analysis to provide a thorough basis upon which to build a theory of change, there is little explanation of how these macro contextual issues will practically translate into ministerial or departmental theories of change. There is a need, therefore, for capacity development, mentorship, stewardship, and technical support to develop the skills and competencies for this kind of work.

There also needs to be a more deliberate process on how to build the evidence and theories for confirming the various hypotheses around how change might happen, and how sectors, provinces, sector departments and programmes, projects, and operational units contribute to these changes through their policy and programmatic choices. These would need to be subjected to evaluation in order to better inform policy choices (for example, around addressing inequality, unemployment, or poverty). As discussed previously, the mechanisms of change in theories of change are often hidden or assumed. Building the evidence for developing explicit, robust theories of change for major policy decisions and programmatic choices may have a significant influence on the determination of outcomes, outputs, and activities at the microcosm of government service delivery and ultimately the successful achievement of national and global development goals.

Conclusion

The past two decades have witnessed a push for recognition of the importance of evidence-informed decision making. It is no longer difficult to convince public policymakers, leaders, and public officials that there is a need for better evidence for decision making, and many governments across the African

continent have championed the institutionalisation of M&E systems. This has led to the normalisation of the use of results-based M&E and particularly the adoption of theories of change in planning processes. One clear challenge that has emerged has been the blind adoption of theories of change as a planning and M&E tool, without sufficient adaptation to the macro planning requirements of the state. Theories of change are very often weak in their fidelity, offering a semblance of a pathway of change that is incomplete and of limited use for planning or measurement. This instrumental use of results-based planning tools such as the theory of change—without due consideration for the depth of reflection, theorising, hypothesising, and testing required in building robust pathways of change—could be addressed through building a transformative consciousness around complexity in development and embracing a systems perspective around the design, implementation, and evaluation of development results.

Governments in Africa need to ensure that deliberate reflection is undertaken to consider the level at which theories of change make sense for planning, prioritisation, and evaluation. Macro-level theories of change should not pretend to be what they are not. The existence of a theory of change at the macro level of planning systems, with all its complexity, should not preclude the need for theories of change at programme and project level. Where theories of change are used—particularly at the programme level—much more attention needs to be paid to causal pathways and mechanisms of change, as well as the need to test these assumptions around how change happens. Theories of change (and their associated implementation frameworks) play a critical role in leading governments to either hit or miss the mark of transformative public-sector delivery—and therefore the powerful role they play in the achievement of development outcomes should not be underestimated.

References

Bakken, L. L., Nunez, J., & Couture, C. (2015). A course model for building evaluation capacity through a university–community partnership. *American Journal of Evaluation*, *35*(4). https://doi.org/10.1177/1098214014523671

Basheka, B. C. (2016). Evaluation capacity building (ECB) in Uganda: Trends, approaches, actors, and the future. *Administratio Publica*, *24*(2), 95–121.

Cloete, F., Rabie, B., & De Coning, C. (Eds). (2014). *Evaluation management in South Africa and Africa*. African Sun Media. https://doi.org/10.18820/9781920689513

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. John Wiley & Sons.

Kusek, J., & Rist, R. (2004). *Ten steps to a results-based monitoring and evaluation system*. World Bank.

McConnell, J. (2019). Adoption for adaptation: A theory-based approach for monitoring a complex policy initiative. *Evaluation and Program Planning*, *73*(April), 214–223. https://doi.org/10.1016/j.evalprogplan.2019.01.008

Murphy, R. J. A., & Jones, P. (2021). Towards systemic theories of change: High-leverage strategies for managing wicked problems. *Design Management Journal*, *16*(1), 49–65. https://doi.org/10.1111/dmj.12068

Rogers, P.J., & Weiss, C.H. (2007), Theory-based evaluation: Reflections ten years on: Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 2007, 63–81. https://doi.org/10.1002/ev.225

SA DBE (South African Department of Basic Education). (2020). *Strategic plan 2020– 2024*. March. SA DBE.

SA DPME (South African Department of Planning, Monitoring and Evaluation). (2019). Guidelines for implementation of the revised framework for strategic plans and annual performance plans. Retrieved from www.treasury.gov.za/?aspxerrorpath=/legislation /pfma/TreasuryInstruction/Guidelines%20for%20Implementation%20of%20the %20Revised%20Framework%20for%20Strategic%20Plans%20and%20Annual %20Performance%20Plans.

SA DPME (South African Department of Planning, Monitoring and Evaluation). (2021). *Revised framework for strategic plans and annual performance plans.* Retrieved from www.treasury.gov.za/legislation/pfma/treasuryinstruction/Revised%20Framework %20for%20Strategic%20Plans%20and%20Annual%20Performance%20Plans.pdf

15 Simple heuristics for drafting theories of change

A case of behavioural insights into food waste

Karol Olejniczak and Igor Lyubashenko

Introduction

Evaluation literature strongly emphasises the need for a participatory approach to developing or modifying public interventions (Cousins & Whitmore, 1998). This postulate is well aligned with public policy and public management literature on the co-production of policies and services (Blomkamp, 2018; Ostrom, 1996).

Co-production or a participatory approach means engaging stakeholders and nonexperts into the processes of sense making of a policy problem and then designing and delivering the solutions that aim at a positive change in the problem situation (King, 2005; Nabatchi et al., 2017).

Our essay focuses on one of the stages of this process—the development of the theory of change, which is simply speaking 'stories that people tell about how problems arise and how they can be solved' (Weiss, 1995). Following the participatory paradigm, we aimed to involve laymen in developing these 'stories'.

The advantages of designing policies and programmes in a participatory way with stakeholders and citizens are straightforward (Bundi & Pattyn, 2022; Cousins, 2003; Smith & Ingram, 2002). First, the participatory approach can provide a richer understanding of the policy issue by including various perspectives and values, making designers more sensitive to certain trade-offs that emerge. Second, it allows tapping into the creativity and local experiential knowledge of the policy actors and generating more innovative solutions. Third, the inclusion of various stakeholders can build a sense of ownership and consensus around policy solutions, facilitating implementation and strengthening the legitimacy of the policy.

However, co-design of public interventions with non-experts can be challenging in two major aspects: user-friendliness and quality of outcomes. The design process has to be user-friendly for people with limited or no prior experience in policy design and evaluation. The concepts have to be simple and intuitive to decrease cognitive load and allow participants to focus on the policy merit. Frames used for designing should also be flexible to fit different policy issues in scale, complexity, sectoral perspective, and the dynamic among actors. Also, the process itself should be arranged to keep the high levels of engagement among participants.

The quality of the outcome—the developed theory of change—is a second challenge. Creative co-design sessions need to end up with well-articulated theories of change. Such a clear outcome has two functions. On the one hand, it provides enlightenment for the participants—at the end of the process, they gain better insights on a specific policy issue. On the other hand, a well-drafted theory of change increases the chances of effective implementation. Therefore, it is worth pondering what approach could facilitate effective participatory work with nonexperts on drafting theories of change for various policy issues.

In this essay, we present an approach to drafting theories of change based on a set of simple heuristics. We argue that the approach could be valuable in ensuring user-friendliness for people participating in the design and in providing the quality of the theory of change. We illustrate our point with a case study involving the test application of the approach with a group of 143 nonexperts who participated in drafting theories of change for the food waste problem in the Warsaw metropolis.

In the remainder of this essay, we explain our approach and provide details on the main steps and concepts used. Then we report on findings from testing the approach with a group of nonexperts. Finally, we close by discussing lessons learned and broader implications for evaluation practice in using theories of change in a participatory manner.

Simple heuristics for drafting theory of change

Conceptual background of the approach

We called our approach a *policy lab*, which is an adaptation of an earlier approach developed and tested with professional policy designers (see Olejniczak et al., 2020).

For our policy lab, we defined the theory of change as follows: *a set of assumptions about (i) groups of actors whose behaviours contribute to the problem, (ii) the causes of those problem behaviours, and (iii) a strategy to induce positive change*.

This definition indicates three specific issues addressed in our approach. First, the approach recognises that public policy follows the basic logic of human problem solving as hypotheses testing. The word *hypotheses* is not used here in scientific terms but in a commonsense way, as simple heuristics used in everyday thinking—assumptions about the problem and its solutions (Evans, 2017). Simple questions can help articulate and compare those assumptions.

Second, the approach concentrates on the behaviours of policy actors. It starts with key system agents and then explores the context that enables their misbehaviours. Establishing who misbehaves, when, and why allows well-focused analysis of the context in which actors make decisions (so-called decision situations), and constructive discussions on determinants of actions taken, particularly the degree of actors' agency. The value of an actor-based behavioural perspective

has been recently recognised in the evaluation of complex programmes (Koleros et al., 2018).

Third, the approach modifies the change logic typical in realist evaluation: context–mechanism–outcome (Pawson & Tilley, 1997). Following suggestions made by Lemire et al. (2020), we distinguish between intervention and the change mechanism. The policy tools are triggers for a mechanism that leads to change. An effective intervention should tinker with a configuration of mechanism elements—neutralising barriers and obstacles to good behaviours, dismantling causes of bad behaviours, and leveraging drivers of good behaviours (context + intervention → mechanism → outcome).

Practicalities of the approach

Our approach in practice consists of two templates—a roadmap of the process and a template to run design sessions. The roadmap of the design process (see Annex at the end of the essay, Figure 15.1) was translated into a specific online working space in Miro, with a canvas covering each step of the design process (see Annex, Figure 15.2). The participants in the processes used post-it notes to fill in the templates with their creative ideas.

The process consists of three stages, each with guiding questions and key terms used for analysis. In order to increase uptake for nonexperts, we used a metaphor of detective work to describe activities of theory of change design. Metaphors are a powerful yet simple way of communicating complex ideas (Morgan, 2006).

In STAGE 1, participants frame the policy issue in terms of behaviours. They need to designate 'suspects' who contribute to the policy problem with their misbehaviour (target groups). Participants develop initial group profiling and can even refocus on a subgroup. They also provide convincing premises for targeting the specific group. This step builds on public policy literature on policy targets (Howlett, 2019) and policy analysis literature on argumentation (Dunn, 2017).

In STAGE 2, participants identify moments of 'crime' (misbehaviour) and speculate about its causes. Their detective work is focused first on reconstructing situations when misbehaviour occurs (action situation) or a chain of events that results in problem situations (decision journeys). Then, once they establish basic facts, they hypothesise underlying causes that drive misbehaviour or hamper positive behaviour along the 'suspect's' decision journey. Finally, they analyse the causes of behaviours in terms of capacity, motives, and opportunity of the target group to behave or misbehave (COM-B). The concept of 'action situations' comes from institutional analysis literature (Ostrom, 2005), decision journeys comes from service design practice (Kalbach, 2016), while the COM-B model is the work of Michie et al. (2011, a nice link to 'means–motive–opportunity'—the classic detective heuristic from Agatha Christie stories).

In STAGE 3, participants look for ways to prevent misbehaviour and replace it with good behaviour. They would think about removing obstacles

to good behaviour and finding leverage points for desired behaviour. At their disposal, they have a spectrum of policy tools. Interventions are organised in a toolbox according to mechanisms that enable behaviours: equip, ban, dis/incentivise, persuade, inform, boost, and nudge. The logic of change comes from behavioural insights literature (Soman, 2017; Weaver, 2015), while the toolbox of policy intervention is inspired by policy tools literature (Bemelmans-Videc et al., 1998; Grune-Yanoff & Hertwig, 2016; Schneider & Ingram, 1990).

Summing up, the proposed process is built in three steps: (1) frame the problem and focus attention, (2) develop a shared descriptive mental model (how things do not work), and (3) devise a prescriptive aspect of what should be done to improve the situation. This three-step approach uses chunking to introduce new knowledge gradually. It also prevents nonexperts from jumping to conclusions too early or—putting it in the words of classic policy analysis authors—enabling 'the art of not knowing too early which type of solution to apply' (Rittel & Webber, 1973, p. 164).

Findings from the empirical testing

The arrangements of policy lab

We tested our approach empirically over five weeks with a population of 143 students working in 41 teams. It was part of the compulsory courses for management students who had no prior experience in public policy and programme evaluation. We called this exercise a *policy lab*.

In the policy lab, students took the role of policy advisors who supported the city of Warsaw during the initial policy design. Their goal was to propose ideas for solutions to reduce food waste in the Warsaw metropolis.

Students worked in 3–4 people teams and participated in 3–5 online sessions (140 minutes each), using Google Meet for communication and the Miro platform with special templates for creative work. Each session followed the same structure: (1) the facilitator delivers a mini-lecture introducing all teams to key concepts and key products of the specific design stage, (2) the teams work in their breakout rooms and use the Miro platform, (3) the participants reconvene to present and compare ideas among the teams and receive initial feedback and clarification from the facilitator, and (4) the teams go back to their creative rooms to improve their ideas.

The final product delivered by each team at the end of the semester was a short policy brief addressed to the authorities of Warsaw. The structure of the brief mirrored three stages of the design process (Figure 15.1). It explained the target group and why it was chosen, diagnosed the situations and causes of misbehaviour, and suggested possible solutions to trigger the appropriate causal mechanism to bring about the desired change. The facilitator graded only the final content of policy briefs.

The findings from the empirical test

We collected various data on both processes and their outcomes. In our analysis, we used three sources: (1) students' individual knowledge retention tested in a quiz (n = 143), (2) students' individual reflection on positives and negatives of the policy lab experience (n = 141), and (3) scoring of final policy briefs obtained by teams (n = 41). We discuss those results below.

The final exam contained—among others—four multiple-choice questions aimed at testing the understanding of key concepts, on which the policy lab was organised. We treated these questions as the basic measurement of students' success in achieving learning outcomes in terms of knowledge.[1] Table 15.1 shows the percentage of correct answers provided by students participating in the course. As shown, the overwhelming majority of students coped well with the task.

During the final exam students answered the following open-ended question: 'In the second part of the semester we worked in a policy lab. We were designing policy solutions on food waste. Please reflect on (1) positives and (2) negatives of your experience with this type of work method. Please provide your opinion in points, clearly explaining your judgement'. Students' answers were coded in three cycles of coding with the use of Maxqda 2022 software. Three groups of comments emerged from the first cycle of coding: (1) on positives (396 of coded segments), (2) on negatives (183 of coded segments), and (3) on main takeaways beyond current course (91 coded segments). The second and third coding cycles identified detailed topics within those groups as described in Table 15.2.

According to participants, breaking down the method to develop the theory of change in three steps was the major positive. If we add to this item other remarks on the lab method and concepts (behavioural insights, routine of work), we see that the lab approach was a dominant positive issue (43.7% of positive codes).

Table 15.1 Results of analysis: Questions that tested students' knowledge

	Percentage of correct answers
Understanding of the model of behaviours	95.10%
Understanding of the public policy design toolbox	88.11%
Understanding of the theory of change	83.92%
Understanding of the three-step procedure of policy lab	89.51%

Source: Authors' own work
n = 143 students

1 We are aware of limitations associated with tests based on multiple-choice questions. This type of exam was selected based on the official regulations of the university as well as considerations of efficiency, taking into account available time resources given to particular courses.

The second major positive was Miro as an online platform allowing for creative work and visual management of ideas (19.2%). It complements the participants' praise for working in teams (18.9%).

On the negative side, the main issue was a specific aspect of teamwork—that is, 26.6% of people did not like that groups were allowed to see each other's work during the process or even intervene with other work. The second negative issue was limited to background research—students wanted more in-depth data on the policy problem and the behaviours of target populations. In real life, this is addressed by desk research and field studies executed between workshop sessions. Some students also pointed out challenging moments in the group process, such as making decisions on the target group, prioritising causes of problems, or choosing among their creative ideas. These problems were reported by a clear minority of policy lab participants. Although these issues are important from the perspective of the policy lab, they may be a side effect of the online setting, in which the lab was conducted.

Finally, two main takeaways for students were the importance of food waste (the exercise nudged them to pay attention to their own behaviours) and the

Table 15.2 Results of analysis: Question on positives and negatives of lab experience

Positives	396 coded segments	Negatives	183 coded segments	Takeaways beyond the course	91 coded segments
Lab method and concepts: three-step procedure	22.7%	Teamwork/ communication	26.8%	Awareness rising	37.4%
Lab method and concepts: behavioural insights	7.6%	Background research	14.8%	Lab method and concepts	23.1%
Lab method and concepts: routine of work	8.3%	Challenging moments	14.2%	Tool—Miro platform	14.3%
Lab method and concepts: general remarks	5.1%	Lab method and concepts	10.4%	Other	25.3%
Tool—Miro platform	19.2%	Tool—Miro platform	9.3%		
Teamwork/ communication	18.9%	Time issue	6.6%		
Topic—food waste	10.6%	Verification of ideas	5.5%		
Other	7.6%	Other	12.6%		

Source: The authors' own work
n = 141 students

potential applicability of the lab method for dealing with other policy or business problems.

The last source of data was students' ideas on policy solutions presented in policy briefs. The course instructor (an academic with practical experience in policy design and evaluation) assessed the quality of this work. Table 15.3 presents the assessment criteria and the average scoring obtained by all 41 teams.

It is worth briefly discussing the actual merit of students' ideas. Teams identified a divergent spectrum of policy targets groups. The most common targets were families with kids and young people running individual households. However, some teams targeted organisational actors, such as restaurants, bakeries, school canteens, and vegetable stands. That divergence indicates that nonexperts were able to grasp the nuanced and complex perspective of the policy problem.

Naturally, that spectrum of different target choices resulted in different problem analyses—from individual shopping situations and cooking routines to ineffective elements of production processes or organisational arrangements. This was followed by a broad spectrum of ideas for policy solutions. The positive was that designers thought in small and practical terms—focusing on simplifying policy addressees' choices, decreasing consumer temptations in specific moments, applying digital solutions to routine processes, and using social

Table 15.3 Scoring the policy briefs

Section of policy brief	Assessment criteria used by the instructor to grade policy briefs of each team	Average score for 41 teams
Framing: choosing target group	Strength of premises for choosing target group From 0 to 5 points where: 5 = clear premises backed up with evidence 0 = unclear premises, no evidence	4.20
Hypothesising about Obstructing Problem (HOP)	Depth of diagnosis From 0 to 5 points where: 5 = in-depth diagnosis with clear plausible hypotheses on causes of misbehaviour 0 = superficial diagnosis, no clear hypotheses on causes of misbehaviour	4.17
Hypothesising about Intervention Type (HIT— ideas for solutions)	Completeness of change mechanism From 0 to 5 points where: 5 = logical causal chain with key change drivers 0 = unclear logic of causal chain	4.10
TOTAL	Maximum available points = 15 points	12.46

Source: The authors' own work
n = 41 teams

leverages to nudge introducing simple comparisons to trigger reflection. On the negative, we observed that several teams still assumed education and information as the main change tool. That contrasts with the latest behavioural insights, indicating 'sermons' as ineffective in inducing behavioural change.

Conclusions

We have proposed a process that drafts theory of change in three steps: (1) framing policy problems in terms of the behaviour of specific groups, (2) developing a shared descriptive mental model (when and why things do not work), and then (3) creating prescriptive aspects on what should be done to improve the situation.

The proposed formula of the policy lab proved to be user-friendly. The chunking of the design steps, the metaphor used to communicate the logic of investigation, and heuristics used during creative parts worked well. In addition, participants perceived the process as generally intuitive and manageable, allowing them to engage in the merit of productive problem solving.

Furthermore, our data indicate that the method helped participants understand the process of creating public policy solutions. We managed to induce genuine engagement of nonexperts into the process. This should be regarded as a valuable step towards achieving the goal of a participatory approach to developing and modifying public policy solutions.

In terms of quality of outcome, our conclusions can be only partial. Our assessment of the quality of content is an academic one. The extent to which the proposed solutions can be used in practice should become a question for follow-up research that could involve practitioners as judges of the ideas generated from the process.

Also, it has to be pointed out that the exercise was conducted among peers. Students' equal standing eliminated any major power struggles. However, in real life, the co-design process attracts different groups of policy actors with different policy standing, interests, and resources. Thus, we should assume that participants' existing social hierarchies and power relations may affect the final outcome. It should also be noted that the group processes among policy lab participants may work differently in ordinary, face-to-face settings. These issues are yet another point for further research.

In conclusion, we can state that the proposed method can become an effective tool that can at least substitute classical consultations with citizens. Engaging 'ordinary citizens' in a policy lab can become a generator of ideas created according to the principle of participation.

Furthermore, we believe this approach is promising in handling complex policy issues in at least four ways. First, it focuses on policy actors who have agency for the actual behaviour change. That translates abstract systems into concrete, manageable elements without losing interconnections.

Second, the complexity and depth of the system are revealed gradually. Thus, it allows laymen designers to immerse into the 'nuts and bolts' of the change

mechanism without cognitive overload and without losing how their part is nested in the bigger picture.

Third, the approach seems to at least partially align with issues described in this volume (see Essay 4 by Dahler-Larsen) regarding the need to push for critical thinking about how the change will happen and which steps are perceived necessary for outcomes.

Finally, when working with numerous teams, the policy lab covers well the multifaceted nature of the problem under consideration. In our case, student teams that spontaneously chose different focuses eventually combined efforts and all aspects of food waste were covered well.

We hope that our case of a policy lab that links evaluation practice with an emerging field of applied behavioural sciences will contribute to an ongoing discussion on improving the engagement of citizens in designing change that matters.

Funding information

References

Bemelmans-Videc, M.-L., Rist, R., & Vedung, E. (Eds.). (1998). *Carrots, sticks & sermons: Policy instruments and their evaluation*. Transaction Publishers.

Blomkamp, E. (2018). The promise of co-design for public policy. *Australian Journal of Public Administration*, *77*(4), 729–743. doi:10.1111/1467-8500.12310

Bundi, P., & Pattyn, V. (2022). Citizens and evaluation: A review of evaluation models. *American Journal of Evaluation*, *44*(4), 574–603.

Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation* (Vol. 9, Kluwer International Handbooks of Education). Kluwer Academic.

Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, *80*, 5–23.

Dunn, W. N. (2017). *Public policy analysis: An integrated approach* (6th ed.). Routledge.

Evans, J. (2017). *Thinking and reasoning: A very short introduction*. Oxford University Press.

Grune-Yanoff, T., & Hertwig, R. (2016). Nudge versus boost: How coherent are policy and theory? *Minds & Machines*, *26*(1–2), 149–183.

Howlett, M. (2019). Behavioural considerations of public policy: Matching policy tools and their targets. In H. Straßheim & S. Beck (Eds), *Handbook of behavioural change and public policy* (pp. 78–88). Edward Elgar Publishing.

Kalbach, J. (2016). *Mapping experiences: A guide to creating value through journeys, blueprints, and diagrams*. O'Reilly.

King, J. A. (2005). Participatory evaluation. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 291–294). Sage. www.loc.gov/catdir/toc/ecip0417/2004009988.html

Koleros, A., Mulkerne, S., Oldenbeuving, M., & Stein, D. (2018). The actor-based change framework: A pragmatic approach to developing program theory for interventions in complex systems. *American Journal of Evaluation*, *41*(1), 34–53. https://doi.org/10.1177/1098214018786462

Lemire, S., Kwako, A., Steffen, B. N., Christie, C., Donaldson, S., & Leeuw, F. (2020). What is this thing called a mechanism? Findings from a review of realist evaluations. *New Directions for Evaluation*, *167*, 73–86.

Michie, S., van Stralen, M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*(42), 1–11.

Morgan, G. (2006). *Images of organization* (updated ed.). Sage.

Nabatchi, T., Sancino, A., & Sicilia, M. (2017). Varieties of participation in public services: The who, when, and what of coproduction. *Public Administration Review*, *77*(5), 766–776. https://doi.org/10.1111/puar.12465

Olejniczak, K., Śliwowski, P., & Leeuw, F. (2020). Comparing behavioral assumptions of policy tools: Framework for policy designers. *Journal of Comparative Policy Analysis*, *22*(6), 498–520. https://doi.org/10.1080/13876988.2020.1808465

Ostrom, E. (1996). Crossing the great divide: Coproduction, synergy, and development. *World Development*, *24*(6), 1073–1087.

Ostrom, E. (2005). *Understanding institutional diversity*. Princeton University Press.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. www.jstor.org/stable/4531523

Schneider, A., & Ingram, H. (1990). Behavioral assumptions of policy tools. *The Journal of Politics*, *52*(2), 510–529.

Smith, S. R., & Ingram, H. M. (2002). Rethinking policy analysis: Citizens, community, and the restructuring of public services. *The Good Society*, *11*(1), 55–60. https://doi.org/10.1353/gso.2002.0017

Soman, D. (2017). *The last mile: Creating social and economic value from behavioral insights*. University of Toronto Press.

Weaver, K. (2015). Getting people to behave: Research lessons for policy makers. *Public Administration Review*, *75*(6), 806–816. doi:10.1111/puar.12412

Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. In J. P. Connell, A. C. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives: Concepts, methods, and contexts* (pp. 65–92). Aspen Institute. www.aspenroundtable.org/vol1/index.htm

Annex: Templates used in lab

Policy Lab for Designing Solutions

Role and goal of the participants

YOU ARE a team of experts who will support The City of Warsaw during their initial policy design

GOAL: propose ideas for solutions aiming at the reduction of food waste in the City of Warsaw. Solution should focus on specific target group

PRODUCT: 3-4 pages of policy brief

3-STAGES of Design Process

(1) FRAMING PROBLEM
Choosing target group

1 Who waste food?

2 What is the scale of the food waste for each of the groups? (Credible sources)

3 What group do we want to focus on? Why?

(2) HOP: ANALYSIS OF THE PROBLEM
Hypothesizing about Obstructing Problem

1 In what situations does food waste take place? (Typical action situations, decision journey)

2 Can we specify misbehavior?

3 What are the causes of misbehavior? (Hypotheses about Problem)

Opportunity: infrastructure, rules, environment

Capacity: has, knows, understands

Motive: feels, calculates, believes

(3) HIT: IDEAS FOR SOLUTIONS
Hypothesizing about Intervention Type

1 What good behavior do we expect?

2 What is already here to enable good behavior?

3 How will the change of the behavior take place? (change mechanism)

Equip

Ban or Restrict

Dis/Incentivize

Persuade

Inform

Boost

Nudge

Figure 15.1 Roadmap for design process

Source: The authors' own work
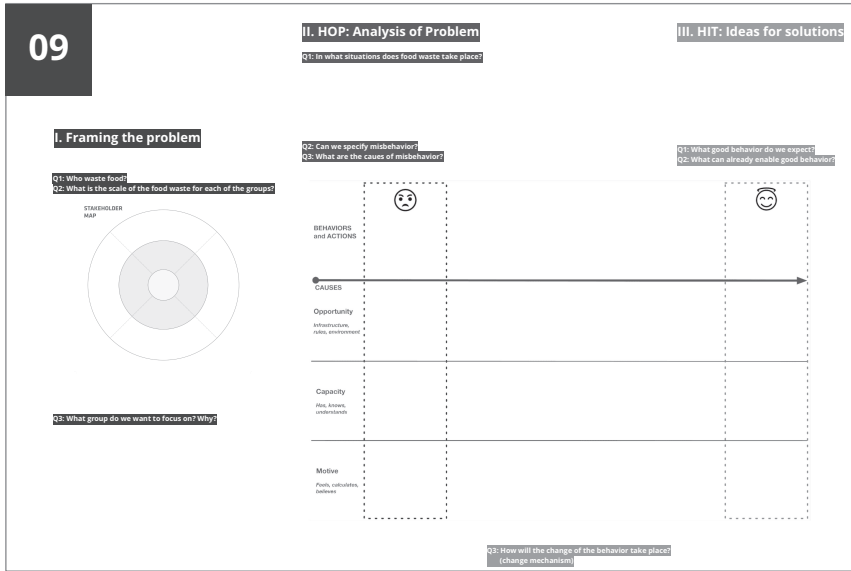
Figure 15.2 Miro design canvas

Source: The authors' own work

Part 5

Applying theories of change approaches for multiple purposes

16 The law of the instrument

Would you rather be a theory or a nail?

Gordon Freer

Introduction

Theories of change are everywhere; they pop up in programme designs, in reviews, in evaluations, and as strategic tools. Several contributions in this volume speak to the myriad ways in which this tool can be utilised, and yet its functionality is not fully exploited.

As an evaluation device, theories of change are discussed widely as useful instruments to better understand underlying programme mechanisms (Chen & Rossi, 1983). It is *de rigueur* in current calls for programme proposals to require a theory of change or for evaluations to review an existing theory of change. They may form part of a briefing for the design of a new intervention, with their mandatory but sometimes misguided inclusion treated as a panacea. But commissioners, designers, and implementers of programmes may be less familiar with the wider application of theories of change as a tool to assist in the design of new (Breuer & Lee, 2015) and the implementation of existing (Freer & Lemire, 2019) programmes, perhaps being more familiar with their use as yardsticks for the evaluation of existing programmes. This lack of familiarity means that theories of change are not fully utilised as tools in design, in implementation, or in reviews and reporting. Instead, to varying degrees, the theory of change is relegated to fulfil a proposal submission requirement or simply reduced to something for evaluators to use—a niche, specialised instrument. The benefits of the tool to assess and comment on the process and mechanisms of change are disregarded. Those less familiar with a project's history forget about the depiction of change in the original programme design documents, perhaps resurrecting it briefly for an evaluation or review (Prinsen & Nijhof, 2015) before quickly moving to assess other programme factors.

The lack of integration or incorporation of a theory of change into the programme design, rollout, management, and review leads the author to query the actual practical purpose of the theory of change. While its theoretical purpose—determining the causes and pathways of change and contributing these to programme design, implementation, or evaluation—remains valid, the absence of its actual wider application raises the question of whether a theory of change has any real-world worth beyond the mind of evaluators. In more optimistic terms,

what is needed for theories of change to have real-world worth beyond the mind of evaluators?

The law of the instrument

Cognitive bias to promote the use of the tools at hand is summarised by Maslow's law of the instrument (Maslow, 1966), which stated concisely is: If the only tool you have is a hammer, then everything is treated as a nail. A corollary of this theory might be that if your understanding of a hammer is limited, you might think it can only be used for hammering nails and not also extracting them, or that all hammers perform the same function—ignoring the nuanced use of claw hammers, geologist's hammers, and sledgehammers. The author argues that the theory of change has fallen victim to Maslow's law, being identified as 'just another hammer', limiting its application to one function and not understanding that it can serve different purposes—a multiuse hammer, to stretch the analogy.

Use of the theory of change may be limited to a component of a business case, or a proposal, or to satisfy the demands of a colleague who has a leaning towards evaluation. Perhaps its inclusion simply serves as a checkbox within a bureaucratic process. This pervasive, unquestioning inclusion has resulted in the instrument being added to a programme's toolbox, without considered reflection on its range of purposes or theory of change thinking across a continuum (Stein & Valters, 2012). Many programme decision makers mistakenly think that it performs the same function as the more familiar tools for programme design and reporting. Instead of exploring the variety of uses of this tool, time pressures, the lack of pressure of the 'immediate', and a focus on contractual deliverables all contribute to the perception that the theory of change lacks practical purpose. This, combined with unfamiliarity with the tool—and an undeserved equivalence with other tools, as 'another hammer'—means the theory of change remains underutilised, mouldering at the bottom of the programme toolbox.

There are many reasons underpinning the lack of greater uptake and use of theories of change, and these reasons are grouped into three clusters, each of which will be explored in more detail: (1) comprehension of purpose, (2) utilisation in programme design, and (3) use in programme implementation. This essay suggests why this tool has fallen victim to these failings and argues that the instrument needs to be seen as an additional, complementary implement that can fulfil several purposes beyond simply satisfying the interests of evaluators, fulfilling bureaucratic processes, or both.

Evaluators then argue that for theories of change to be better leveraged or utilised, certain actions are essential; firstly, in using a theory of change as a crucial part of programme design, and secondly as an assimilated programme instrument for reporting, reflection, strategic decision making, and learning.

Comprehension of purpose

At its most basic, within this first cluster, those seeking to understand a programme are often confronted with an amalgamation between logical frameworks and theories of change. Many are well acquainted with the linear thinking displayed in logical frameworks and assume the same format and purpose for this 'other' instrument, thinking it adds limited value (Mayne, 2017) because of failing to appreciate the need to also take cognisance of complexity. The tendency for linear process fails to consider the need to track the complexity of programme context and its effect on rollout. Echoing the ability of the tool to handle complexity, Tangelder (Essay 20) and Lemire (Essay 24) separately speak to the increasing use of theories of change to contribute to systemic change thinking and to underpin inclusive transformative change. This lack of appreciation of the differences between more linear tools and tools that can better reflect complexity becomes apparent in the absence of management utilisation, which is explored later in this essay. A lack of understanding and appreciation of the different tools' functions and uses contributes to this amalgamative thinking (Freer & Lemire, 2019). Confusion behind the potential value addition and a lack of appreciation or absence of the necessary skills to wield this tool (Maru et al., 2018) adds to its relegation.

Part of the blame for this lack of comprehension may lie at the feet of the evaluation community, who perhaps too rarely clearly explain the variety of uses of a theory of change. Mayne (2017), amongst others, outlines that a theory might be *inter alia* a product on its own, a process to facilitate agreement amongst stakeholders, a design tool, or a foundation for an evaluation. While those familiar with the tool might wield it in a variety of ways to produce unique products that serve different purposes, these wider, deeper applications may not have been effectively communicated to those outside the community. If evaluators use a theory of change at a specific time, or to fulfil a defined purpose without commentary or inclusion, they may well be limiting those outside this clique to understand the rich variety of functionality that this tool offers.

Utilisation in programme design

The second cluster focuses on what is often the first exposure programme staff have to the theory of change—the design process of the theory itself. As a result of several challenges explored in more detail in this section, the theory of change is often not seen as being a practical, contributory component that can assist in programme implementation or reporting if done well at the design stage. Consequently, its utilisation is relegated, further obscuring its purpose and function.

Why might that be the case? Reflecting on Mayne's comments mentioned previously, Perrin (2021, and in this volume) argues that it is because too often theories of change are not fit for the purpose, with some being too simplistic and others watered down to an extent that they are, in essence, irrelevant. At best, these theories are impractical; at worst, they are harmful. Others (James, 2011; Mayne, 2017;

Vogel, 2012) have also pointed to this oversimplification or the vague, generic con-struction of theories of change that add little value to programme design. Aside from displaying complexity, this tendency may be because of a lack of clarity around the limits of the theory's reach (Weiss, 1997), or it may result from the pref-erence towards linear cause-and-effect thinking. In addition to these more ethereal reasons, Tangelder (Essay 20) offers a range of practical reasons (and solutions) why programme staff may fail to engage in the design of a theory of change.

In addition to these reasons, my experiences suggest that they are often regarded as a bureaucratic essential (Rogers, 2007), being demoted in favour of the more visible and, in some cases, contractually central logical framework or logframe. The quantifiable targets and milestones, and the time-limited expected dates of delivery of a reporting framework, are easily understood and can be transposed into a work plan with easily identified key performance indicators. For funders, these can be translated into definite, objective, measurable markers for accountability (Gasper, 2000). In contrast, theories of change focused per-haps on more indistinct behaviour change speak to more obscure indicators—and more troublingly might also show regressive change, displaying a negative programme impact—and to longer-term impact (Prinsen & Nijhof, 2015). While these indicators and pathways of change may tell the story of the goal the programme seeks to contribute towards, often this impact is so far beyond the programme's influence and lifespan that it is beyond the scope and mental horizon of programme staff, who are focused on achieving the next quantifi-able milestone. These longer timeframes echo the limits of reach mentioned by Weiss (1997). While similar arguments were made regarding impact statements in logical frameworks (Gasper, 2000), the incremental, measurable, contributory steps of outputs stand in contrast to the sometimes less tangible transformations identified in theories' pathways to change.

The proposed objective of the programme might be made more remote by a general absence of grounding theories of change in established social science theory (Vaessen & Leeuw, 2010) and by failing to demonstrate these established behaviour change patterns in the theory design process. This further removes the reality of longer-term change from the minds of programme staff, and this, in turn, lessens the importance or centrality of the programme theory.

This longer-term perspective, the scarcity of clear time-based parameters, the possibility of regressive change, and an absence of a sense of the 'immediate' tied to programme performance all contribute to the perception that the pro-gramme theory lacks practical purpose, and it concedes to more practical con-siderations about the design of the programme.

Use in programme implementation

The third cluster of identified reasons concentrate on the use of the theory dur-ing programme implementation. The effective purpose (cluster one) and design

(cluster two) of a theory of change can underpin the extent to which a theory of change is employed and leveraged as a tool by implementers. Freer and Lemire (2019) argue that simplified programme theories limit their usefulness as programme guides and the resultant commitment to their use and application by decision makers:

> When the underlying design and purpose of a theory of change are not well understood, or perhaps even misunderstood, the subsequent use of the ToC—by evaluators to gather data or by programme staff as a tool to inform decisions—is also likely to be misapplied.

As mentioned, programme success is often measured on the achievement of key performance indicators, and all effort is focused on this achievement. If the theory of change is not incorporated into the process of reporting and assessing programme progress, then reporting or assessing against the same is disincentivised.

Depending on the level of involvement and inclusion of programme staff in the design of the programme theory and the extent to which it is incorporated into regular operations and reporting, even an annual review is often a predictable routine rather than an exercise in deepening a programme's understanding of its operational context. When a programme is performing well or overachieving, reporting is often limited to the agreed metrics and no time is taken to clearly understand why a target has been reached or exceeded. Mapping programme activity against its theory of change may give some insight and provide guidance regarding the next strategic steps.

Towards resolution

Together these views—a misunderstanding of its purpose and a lack of its use in programme design and rollout—indicate that this lack of well-understood uses of a theory of change influences its abandonment as an effective programme tool. By disregarding the richness and depth than can resonate from a well-designed and applied theory of change, programme stakeholders lose a more nuanced analysis and understanding of programme purpose, implementation, and impact.

A number of adaptations can be proposed to overcome some of the challenges identified above, with the hope that these initial steps will overcome the current inertia and assist theory utilisation to gather momentum.

Defining purpose

Given the broad range of applications and uses of a theory of change, stakeholders should be specific about the purpose of using a theory of change at a particular point in the programme life cycle. This is more clearly articulated in the various phases of programme conceptualisation, design, implementation,

and review. Using the tool through these phases might be iterative and mutually reinforcing, creating richer, more detailed programme commentary. While not calling for limiting a theory of change's purpose to a single function, such specificity may more effectively communicate the purpose of using a theory of change at that point, rather than allowing for a broader, open interpretation.

This is not a new call. Delahais and Toulemonde (2012) suggest narrowing the focus of theory-based evaluations. Maru et al. (2018) call for clarity of purpose and a recognition of the multiple uses of the tool. Freer and Lemire (2016) call for consideration and the deliberate selection of the theory components in a programme evaluation rather than suggesting the safe but generic and blunt call for a theory-based evaluation. In the same vein, understanding that a theory of change does not make a good programme, and in keeping with Perrin's (2021) thinking, programme commissioners can also opt to exclude a programme theory of change from the various programme phases, if it serves no purpose.

This scope and purpose, including the deliberate omission of a programme theory, should be considered in the overall programme design. The defined scope might be expanded later, but an expansion then becomes a deliberate action, requiring engagement regarding the theory's use and purpose taking into consideration the value it will add to the programme.

Inclusive but applicable design

It is tempting to call for the broad inclusion of stakeholders in the process of designing a theory of change, but this call is often misinterpreted as involving all in the design process. Prinsen and Nijhof (2015) identify inclusion in theory design as a contentious issue, with some promoting including all actors and stakeholders, while others advocate limiting inclusion to specialists. Similar thinking regarding inclusion in the formation of logical frameworks has been met with limited success (Bakewell & Seka, 2005).

While inclusion is often a preferred option, the level of inclusion must be tempered by the defined purpose of the programme theory, recognising the limits and specialities of the programme staff. To stretch the analogy of the toolbox a little more, broad inclusion would equate to requiring all craftsman to perform all tasks, rather than recognising that some craftsmen are more skilled in specific areas. Designing a programme theory should involve communicating the purpose and process of the theory to all programme staff, involving them in the design process to the extent that it adds value to both the design and their role in implementation. Even if their role in the actual design process is limited, the process and purpose of designing the programme theory should be regularly communicated to all staff.

Grounding the theory of change in established social science should recognise the extent to which the programme might contribute to change within its set time

limits. These limits should form a foundational part of the communication to the programme staff, ensuring their understanding that their level of achievement is realistic. This will make the application of the theory of change more applicable, recognising that their day-to-day programme implementation contributes to achievable change. Theories of change tend to focus on longer-term implications—on the 'bigger picture' or complexity of the system the programme operates within—and on the programme's underlying *raison d'être*. It is important that programme staff do not lose sight of this longer-term direction and purpose in the day-to-day operations of implementation (Freer & Lemire, 2019).

Inclusion in programme implementations

With a clearly defined purpose and a distinct and demarcated design process, the possibility exists for the theory of change to act as a contributory programme reporting tool. Programmes could clearly differentiate the purpose and function of the tools it has available and take advantage of the potential symbiotic relationship between them (Freer & Lemire, 2019).

The symbiosis can be leveraged, with the familiar quantifiable logical framework milestones tied to more qualitative indicators of progressive (and possibly regressive) change. If needed, these indicators may act as early warnings to inform programme adaptation or, in the case of programme achievement, to suggest realistic, revised 'stretch' targets (Uwizeyimana, 2020). This cooperative, interdependence of reporting tools would involve a step away from the more common, linear programmatic thinking. Instead of forcing the programme theory to act as a hammer, replicating other reporting and management tools, symbiotic reporting would allow the theory of change to be established as a complementary tool in its own right, adding texture and depth to programme commentary.

In this reporting process, the programme theory is used as a reflective, dynamic tool (Perrin, 2021) that records programme change and, if needed, adaptation, through a regular review reflecting the inherent 'messiness' of development work (Bakewell & Seka, 2005). Over the longer term, this will inform learning for both broader programme design as well as the opportunity to refine the established social science theories on which the programme theory is based.

Concluding thoughts

While the pervasiveness of theories of change in programme design and evaluation may be seen as a positive step, this ubiquitous, uncritical presence is, in fact, more damaging to the true function of the tool. Instead of being viewed as an instrument of precision, its commonplace inclusion has, at best, led to confusion regarding its purpose and a butchering of its use. At worst, it has resulted in this tool being used only by specialists, with its potential use being ignored and programmes not being able to utilise its purpose.

Through the careful reflection of some actions, a more considered inclusion of theories of change is called for, with finely calibrated tools for programme conceptualisation, design, and reporting. Rather than requiring a theory of change to perform a function better suited to another implement, the nature of the theory of change as a tool should be better understood and this tool be employed with the precision that it deserves. Just as the thinking underpinning the logical framework has adapted through several iterations and has been supplemented with training on its correct usage, so too is there a need for greater utilisation of the theory of change by understanding its richness and diversity of application. If not, there is a likelihood that the criticisms levelled at the theory of change from its detractors will echo those levelled at the logical framework decades ago:

> Like every such 'formalised' system it could only too easily degenerate into another piece of bureaucracy if not applied imaginatively and intelligently. (Cracknell, 1989, quoted in Prinsen & Nijhof, 2015)

References

Bakewell, O., & Seka, A. G. (2005). *The use and abuse of the logical framework approach*. Swedish International Development Cooperation Agency.

Breuer, E., & Lee, L. (2015). Using theory of change to design and evaluate public health interventions: A systematic review. *Implementation Science*, *11*(63). https://doi.org/10.1186/s13012-016-0422-6

Chen, H. T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, *7*, 289–302.

Cracknell, B. (1989). Evaluating the effectiveness of the logical framework system in practice. *Project Appraisal*, 4(3), 163–167. https://doi.org/10.1080/02688867.1989.9726727

Delahais, T., & Toulemonde, J. (2012). Applying contribution analysis: Lessons from five years of practice. *Evaluation*, *18*(3), 281–293.

Freer, G., & Lemire, S. (2016). *The keystone node approach: Conducting theory based evaluations of complex programmes* [Conference presentation]. United Kingdom Evaluation Society Conference, London, 27–28 April.

Freer, G., & Lemire, S. (2019). Can't see the wood for the logframe—Integrating logframes and theories of change in development evaluation. *Canadian Journal of Program Evaluation*, 33(3), 336–353.

Gasper, D. (2000). *Logical frameworks: Problems and potentials*. 3-167. Teaching Material for ISS Participants.

James, C. (2011). *Theory of change review: A report commissioned by Comic Relief*. Retrieved from www.actknowledge.org/resources/documents/James_TOC.pdf

Maru, Y. T., Sparrow, A., Butler, J., & Banerjee, O. (2018). Towards appropriate mainstreaming of 'theory of change' approaches into agricultural research for development: Challenges and opportunities. *Agricultural Systems*, *165*, 344–353. https://doi.org/10.1016/j.agsy.2018.04.010

Maslow, A. (1966). *The psychology of science*. Harper & Row.

Mayne, J. (2017). *'Good' theories of change*. The Evaluation Centre for Complex Health.

Perrin, B. (2021). Theories of change: Who needs them? European Evaluation Society Conference. Online, 6–7 September.

Prinsen, G., & Nijhof, S. (2015). Between logframes and theory of change: Reviewing debates and a practical experience. *Development in Practice*, *25*(2), 234–246. https://doi.org/10.1080/09614524.2015.1003532

Rogers, P. (2007). Theory-based evaluation: Reflections ten years on. *New Directions for Evaluation*, *114*, 63–81.

Stein, D., & Valters, C. (2012). *Understanding 'theory of change' in international development: A review of existing knowledge*. London School of Economics.

Uwizeyimana, D. E. (2020). The logframe as a monitoring and evaluation tool for government interventions in a chaotic and complex environment. *Africa's Public Service Delivery and Performance Review*, *8*(1). https://doi.org/10.4102/apsdpr.v8i1.328

Vaessen, J., & Leeuw, F. L. (2010). Interventions as theories: Closing the gap between evaluation and the disciplines. In J. Vaessen, & F. Leeuw (Eds.), *Mind the gap: Perspectives on policy evaluation and the social sciences* (Vol. 16). Transaction Publishers.

Vogel, I. (2012). *Review of the use of theory of change in international development*. UK Department for International Development.

Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review*, *21*(4), 501–524. https://doi.org/10.1177/0193841X9702100405

17 3ie

A 'balloon-squeezing' approach to the theory of change

Marie Moland Gaarder

3ie's mission-level theory of change

Most institutions have an overall mission or objective statement that indicates what it hopes to achieve; this can be interpreted as the overarching theory of change of the institution. In the case of the International Initiative for Impact Evaluation (3ie), this aim was embodied in its mission as stated in the founding document: 'contribute to the fulfillment of aspirations for wellbeing by encouraging the production and use of evidence from rigorous impact evaluations for policy decisions that improve social and economic development programmes in low- and middle-income countries' (3ie, 2008). This was further elaborated on in an introductory note: 'The International Initiative for Impact Evaluation (3ie) seeks to improve the lives of poor people in low- and middle-income countries by providing, and summarizing, evidence of what works, when, why and for how much' (Gaarder & White, 2009). Simply put, the underlying theory of change was to provide funding for effectiveness evidence; this would lead to its production, the evidence produced would be relevant, syntheses and evidence platforms would help improve access, access would increase use, which in turn would result in more effective policies and programmes.

3ie management knew that, for each step in the causal chain, various assumptions would need to hold true. With the benefit of hindsight, I would like to suggest that over the subsequent 15 years the organization has used the balloon-squeezing approach to the theory of change. When you squeeze a balloon, the air gets displaced within the balloon but does not leave it (the balloon would merely bulge elsewhere). If you think about the air in the balloon as the obstacles to using evidence to improve development policies and programmes, then squeezing the balloon is what you do when you test one of the underlying hypotheses of 3ie's theory of change. If the squeeze is unsuccessful or meets some resistance, it would imply that some of the underlying assumptions do not hold true. You would always start the squeezing at the far-left side of the 'theory of change' balloon, namely where the assumptions about inputs and outputs are being made. For anyone wondering why this is the necessary place to start: you cannot test whether something is used and whether it has the desired effect (that is, the right-hand side of the theory of change), if it has not been provided in the first place.

To put the balloon to the test, you need to add hands and squeeze the balloon in more and more places, moving right. This is what is done below. Spoiler alert: We have yet to squeeze sufficiently in enough places to test whether it explodes.

Squeeze One: Funding leads to production

In 2006, few people in the development field had heard about impact evaluations as a useful tool to inform development policy, and few of these studies had been carried out. There was no need to focus on specific evidence gaps, argued a seminal report published by the Center for Global Development (CGD), with the telling title 'When will we ever learn?', as the development field was simply a huge black hole (CGD, 2006).

3ie was launched in 2008, in part as a response to a recommendation made in that CGD report. It was generously funded by the Gates Foundation, the Hewlett Foundation, and the Department for International Development (now the Foreign, Commonwealth & Development Office). We launched well-funded and widely publicized evidence windows and assumed that the supply side would respond. Indeed, it did and from the first open window, which received 78 proposals, the subsequent windows received several hundred.

3ie, as a large promoter, manager, and funder of rigorous impact evaluations globally was a major contributor to the impact evaluation revolution that followed, alongside universities (for example, Berkeley), and organizations (such as the Abdul Latif Jameel Poverty Action Lab and Innovations for Poverty Action). Researchers were getting more traction in the development community for the argument that they could not know whether programmes and policies contributed to the UN Millennium Development Goals without impact evaluations; a robust counterfactual was needed.

Fifteen years on from the CGD report, the body of evidence has exploded. From only a few hundred impact evaluations, 3ie's Development Evidence Portal[1] now includes more than 10,000 studies (see Figure 17.1). The number of institutions that can carry out impact evaluations has also multiplied, thus reducing any production bottleneck to counterfactual evidence.

Squeeze One has been successful, although gaps still exist in many sectors. To a first approximation, we would expect the investment in generating evidence about a sector to match up with investment in programming in that sector. That is, the three bars for each sector in Figure 17.1 (representing the percentage of impact evaluations, systematic reviews, and official development assistance (ODA) in that sector) should be roughly equal in length. When a sector draws a substantial proportion of aid spending, this should be justified with a broad evidence base about what works in that sector. But this is not necessarily what we observe. For example, while the social sectors (health, nutrition, and education), which receive about 28% of development assistance globally have a 55% share of all impact evaluations registered in the Development Evidence
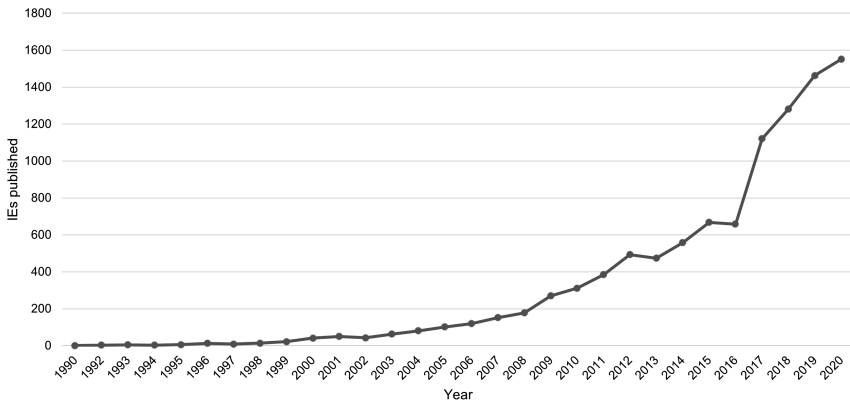
Figure 17.1 Impact evaluation publications on effectiveness of development interventions, by year

Portal, the transport, energy, and governance sectors, which received about 35% of development assistance over the same period, have only 7% of the total of impact evaluations (see Figure 17.2). What is more worrying is that the initial funding enthusiasm appears to be waning. This is worrying not only because of remaining knowledge gaps but because most development interventions do not work as well as they could, and hence continuous evaluation and improvement is needed in all sectors.

Source: Impact evaluation and systematic review data from 3ie's Development Evidence Portal, accessed March 2022 and reproduced with permissions from 3ie; ODA data from stats.oecd.org

Squeeze Two: The evidence produced would be relevant

As the evidence field exploded, various shortcomings of the early crops of impact evaluations also became apparent. We had a front-seat view in 3ie, where we developed a range of screening criteria for good impact evaluations and screened and reviewed hundreds of proposals and final reports. The criteria included:

- Having a clear theory of change
- Focusing on important policy-relevant questions and engaging with stakeholders
- Using a variety of methods to get a good understanding of the underlying causes for impact or lack thereof
- Having diversity in team composition, including members with strong contextual understanding
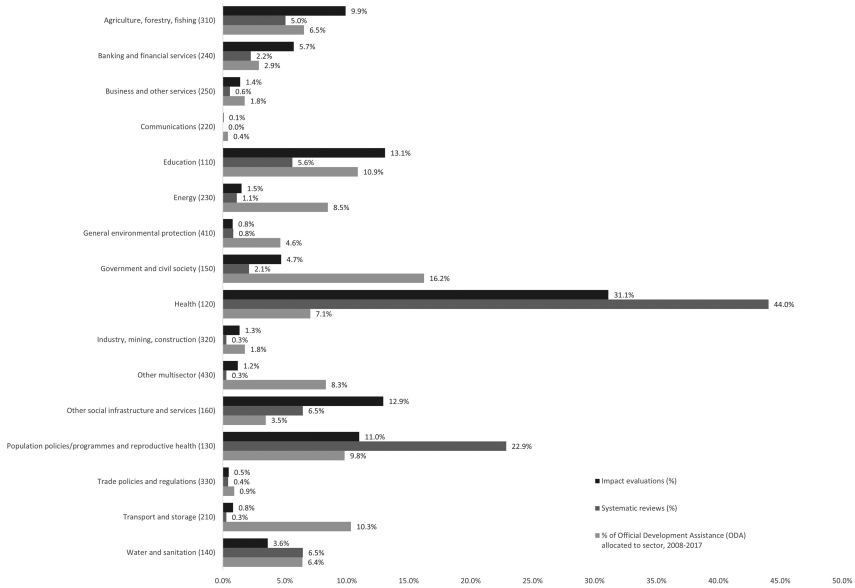
Figure 17.2 Impact evaluation publications by sector compared to development assistance spending

- Including an analysis of cost-effectiveness or cost-benefit
- Exploring equity considerations and the heterogeneity of effects
- Paying sufficient attention to unintended effects and to ethical considerations.

Very few proposals that we received fulfilled even a subset of these criteria, and we used the review process to select the better ones and helped further strengthen the proposals in these areas. Although significant progress has been made over the last decade, the field is not yet where it needs to be in terms of providing policy-relevant impact evaluations. For example, only about one in five impact evaluations includes a good cost-effectiveness analysis, thus missing the opportunity to respond to the questions often foremost on policymakers' minds (Brown & Tanner, 2019). The incorporation of equity considerations into impact evaluation and systematic reviews is also seriously lagging (see Figure 17.3).

Therefore, Squeeze Two has to date only been partially successful. A number of underlying assumptions do not sufficiently hold true, including but not limited to the following: (1) all funders of impact evaluation are able to discern high-relevance impact evaluations from the rest and are willing to finance these; (2) all researchers are interested in the policy relevance of their work; (3) there is general agreement on what constitutes a high-quality, policy-relevant
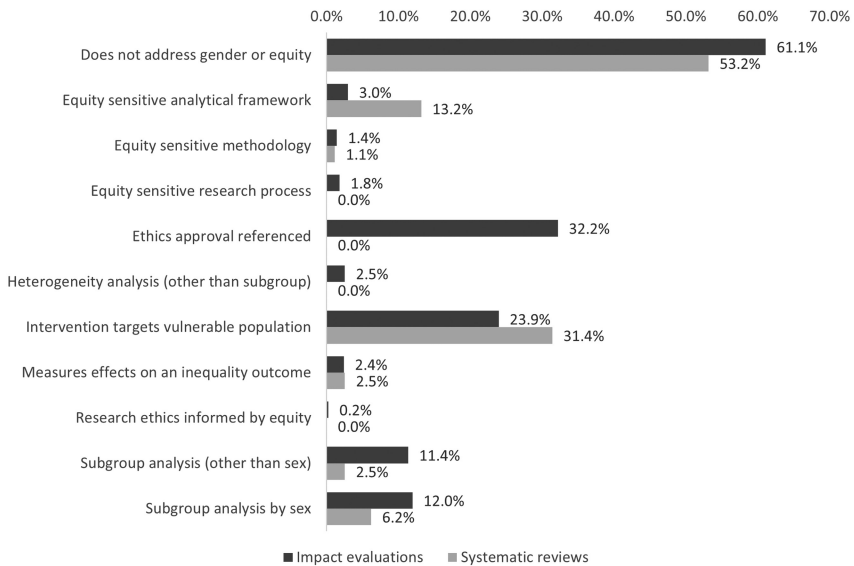
Figure 17.3 Equity considerations in impact evaluations and systematic reviews

impact evaluation among impact evaluation practitioners; and (4) we all know and agree on the exact trade-offs that exist when making decisions about the design and implementation of an impact evaluation (for example, dropping subgroup analysis from an impact evaluation will be cheaper and faster and may hence inform policy faster, but it will not contain crucial information about how the programme affected some particularly vulnerable groups, possible adversely).

Having realized this, 3ie continues to play an important role in convening conversations, promoting improvements, and working with partners to improve guidance and training in topics like cost-effectiveness; equity considerations; process evaluations; transparency, reproducibility, and ethics; stakeholder engagement; and contribution tracing.

Squeeze Three: Syntheses and evidence platforms improve access

One problem with the exploding field of impact evaluations was that unbiased access to the best evidence became a challenge. Decision makers did not have the time to read all the relevant studies, nor in most cases did they have the requisite skills and information to know where to find it and which studies to trust. On top of that, most of the studies are written in a manner that is actionable and accessible only to researchers.

To address these challenges, 3ie undertook a number of measures over the years. First, 3ie took the lead in creating a systematic review group within the Campbell Collaboration. Systematic reviews ensure a systematic and unbiased process for summarizing the body of evidence on a specific topic. In 2010, the International Development Coordination Group was created with 3ie hosting the secretariat and becoming a lead producer of systematic reviews in the development field. Second, 3ie developed the Evidence Gap Map tool, which is influenced by a theory-of-change, matrixed approach to mapping the intervention and outcome evidence for a certain theme. This provides an overview of both existing evidence as well as the gaps in an area of interest, and it can be used both to get rapid access to evidence as well as insights into priority research areas where more evidence is crucial. In addition, 3ie invested its core resources in developing the aforementioned Development Evidence Portal—the largest searchable database of effectiveness evidence on development in existence. Most recently, we developed a help-desk approach and 'just-in-time' services whereby trained personnel respond to evidence inquiries from policymakers by extracting and summarizing the most relevant evidence available in a plain-language and action-oriented format.

We, and others, are making evidence rapidly available through reviews and syntheses, through help desk-type services for evidence translation and rapid evaluations. We are seeing a marked uptake and interest in such products from the donor and funding side, as well. Hence, the access to curated and translated evidence has increased dramatically, lowering further any barrier to use by decision makers. Nevertheless, the limitations that many impact evaluations suffer from—as discussed under Squeeze Two—also affect the syntheses of this evidence and impose some limits to the relevance of what we are improving access to. An example is that few systematic reviews look at the cost-effectiveness of policies, simply because so few impact evaluations have analyzed this.

Squeeze Four: Access increases use

While weaknesses remain with some of the evidence, access to relevant, practical, and digestible evidence has improved a great deal over the last five years for implementers and policymakers. Even so, this is unlikely to be enough to make continuous use of evidence the default in international development operations. As long as what is traded in the international development marketplace are good intentions, outputs and resources, rather than outcomes and development effectiveness, it will be an uphill battle to turn the focus to evidence-informed decision making (Gaarder & Bartsch, 2015). Several institutional evaluations have assessed how international development institutions like the World Bank, Norwegian Agency for Development Cooperation (Norad), and German Development Cooperation learn and use evidence and what the main drivers and barriers are (World Bank Group, 2014, 2015; Norad, 2012, 2014; German

Institute for Development Evaluation, 2021). While each is a very different institution with differing track records around their in-house research, data production, use, and literacy, the key findings are strikingly similar:

- There is a lack of institutional incentives, consistent signals, and role modeling from the top of the organizations that learning and evidence use are important.
- Quality review and approval meetings often do not add value but just wave things through. There is a mindset of ticking the boxes and making it look good.
- Success is measured by project approvals and disbursements of funds, not by results on the ground, which come at a time when most of those involved have moved on.
- Time and resources for monitoring and evaluation training and learning from evidence are not prioritized.

Other literature looking at evidence use in policy in Africa comes to strikingly similar observations: 'A core message is that evidence use is complex and begins long before an evidence journey starts. Evidence use needs to be planned for and woven into the institutional culture. This needs active facilitation of the process, often in a knowledge brokering role which manages both the supply of and the demand for evidence' (Goldman & Pabari, 2020, p. 224).

With these insights emerging and it becoming more apparent that more evidence was available—but not used or used well—3ie decided to update its theory of change in its strategy 2021–2023.[2] We realized that for access to lead to increased, other factors needed to be addressed. We needed to invest in long-term partnerships with organizations and governments; to invest in the local capacity to produce, translate, and use evidence; and to invest in understanding and improving institutional evidence cultures.

To help address these institutional and structural barriers to improved evidence use, 3ie and other specialized institutions in the evidence space can play an active role. In the case of 3ie, our efforts can be summarized under the headlines of *understanding*; *advocating*; and *collecting, evaluating, and learning*.

When working with specific partners and stakeholders, we are increasingly prioritizing *understanding* how knowledge and learning are mediated through interpersonal exchanges, and how team dynamics and connections to social networks shape the potential for learning and knowledge sharing. We also prioritize understanding what moments in a project or policy cycle are amenable to being evidence-informed and by what type of evidence.

Given the important and consistent finding that learning and knowledge sharing are only likely to flourish if there is senior management commitment, leadership, signaling, and role modeling, 3ie is *advocating* for this increased commitment from the top of international development institutions. This

generally takes the form of joining calls for action, panel discussions, and postings on social media.

Finally, we are committed to *collecting*, *evaluating*, and *learning* about promising incentives, processes, and approaches that reward learning and the use of evidence or at least contribute to removing the existing barriers. In 2022, we started the exercise of collecting promising measures through crowdsourcing, expert inputs, and literature searches. To focus on incentives, we collect ideas on how to improve incentives for staff at all levels to question assumptions, look for and use the best available evidence, and continuously learn (for example, in performance reviews, promotion decisions, and through resource availability). We also explore how to improve institutional incentives, such as by creating joint responsibilities for achieving intended outcomes by international development institutions and partner countries or by rating institutions on their commitment to a strong evidence culture in their internal processes and commitment to adopting promising measures. To focus on processes, we collect experiences on how to best ensure that sufficient root-cause analyses and evidence are built into programmes prior to approval (quality at entry). We also look into the procurement processes that are more conducive to helping inform development programmes with good evidence. Other measures could include, for example, providing trainings on *how* to use *what* type of evidence *when* throughout a project or policy cycle (and *by whom*).

Squeeze Four, our most recent addition to our institutional theory of change, is itself a composite of many subtheories of change. How do you change an institutional culture? How many lessons are generalizable, and what are institution-specific measures? Can outside institutions, especially if they act in unison, influence institutional cultures? How can they best do this—is there a combination of lobbying, informing, and training that works particularly well? We will use our own contribution-tracing tools to evaluate and inform our approaches.

What happens next? The most likely scenario is that the improved understanding we and the wider field will gain of how to improve use is likely to lead to new opportunities and challenges along the theory of change—with new gaps and better approaches constantly being uncovered—and hence the air in the balloon will be further displaced.

Conclusion

As evaluators, we are used to discussing with operational staff the fact that they need to be responsible for and measure uptake and impact of the intervention, not only the provision of services and outputs that are under the program's control. The same argument holds for organizations like 3ie that are committed not only to the production of relevant evidence but also to its good use in decision making so that development programmes become more effective and lives are improved. In this essay, it has been argued that the obstacles to using evidence

to improve development policies and programmes are like the air in a balloon. Over the years, 3ie has tested or squeezed subsequent hypotheses by contributing to the increased production of development evidence, the improved relevance of the evidence, as well as the improved access.

The essay argues that these are necessary but not sufficient conditions to make evidence use the default in international development work if institutional and structural barriers to improved evidence use persist. 3ie has realized that for access to lead to increased use, we need to invest in long-term partnerships with international development organizations and governments; in the local capacity to produce, translate, and use evidence; and in understanding and improving institutional evidence cultures and incentives. We are updating our theory of change accordingly, and are working more directly with governments and donor organisations to facilitate their use of evidence, including through the Global Evidence Commitment launched in October 2023, which brings together major funders in a shared commitment to specific forms of evidence use.

Notes

1 The Development Evidence Portal (https://developmentevidence.3ieimpact.org/) is the world's largest database of impact evaluations and systematic reviews on interventions to improve lives in low- and middle-income countries. It is updated monthly and contains more than 10,000 impact evaluations, systematic reviews, and evidence gap maps. The Development Evidence Portal provides advanced searching capabilities and a rich set of options for filtering studies to help you find the evidence you need.
2 www.3ieimpact.org/about-us/what-drives-us/strategy

References

3ie. (2008). *Founding document for establishing the International Initiative for Impact Evaluation 3ie*. 3ie Founding Document. 3ieimpact.org

Brown, E., & Tanner, J. (2019). *Integrating value for money and impact evaluations: Issues, institutions, and opportunities*. Working Paper No. 9041. World Bank Policy Research.

Center for Global Development (CGD). (2006). *When will we ever learn? Improving lives through impact evaluation*. The Evaluation Gap Working Group.

Gaarder, M., & Bartsch, U. (2015). Creating a market for outcomes: Shopping for solutions. *Journal of Development Effectiveness*, *7*(3), 304–316.

Gaarder, M., & White, H. (2009). The International Initiative for Impact Evaluation (3ie): An introduction. *Journal of Development Effectiveness*, *1*(3), 378–386.

German Institute for Development Evaluation. (2021). *Rigorous impact evaluation: Evidence generation and take-up in German development cooperation*. Research Report.

Goldman, I., & Pabari, M. (Eds.). (2020). *Using evidence in policy and practice: Lessons from Africa* (1st ed.). Routledge. https://doi.org/10.4324/9781003007043

Norad. (2012). *Use of evaluations in the Norwegian Development Cooperation system*. Evaluation study.

Norad. (2014). *Can we demonstrate the difference that Norwegian aid makes?* Evaluation report.

World Bank Group. (2014). *Learning and results in World Bank operations: How the bank learns. Evaluation I*. Independent Evaluation Group, the World Bank.

World Bank Group. (2015). *Learning and results in World Bank operations: Toward a new learning strategy. Evaluation II*. Independent Evaluation Group, the World Bank.

18 Integrating theories of change in programme management and delivery

Mark Oldenbeuving

Introduction

From its origins in evaluation, theories of change have become common practice for the design and implementation of development programmes (Stein & Valters, 2012; Valters, 2015). Most donors now expect a theory of change—as a product—to be developed or updated as part of the start-up phase of a programme, and there are numerous guides available for the development of theories of change (Funnell & Rogers, 2011; Vogel, 2012).

As others in this book have argued, the process of *developing* the theories of change is equally, if not more, important. The ways in which theories of change get developed as part of programme design and start-up vary greatly, but a common scenario might be one where the leadership of a programme brings together key staff and stakeholders in a participatory process to co-design the theory of change based on the donor requirements. This process could be led by the programme's monitoring and evaluation team, who host a set of working groups and bring in an external consultant to lead a theory of change design workshop with the programme team, donor representatives, and several external stakeholders. Following the workshop, a draft theory of change is circulated for feedback and visually improved by a graphic design consultant. If done correctly, this process is likely to result in a robust theory of change that the programme team buys into and is enthusiastic about incorporating into programme activities.

However, even in the above scenario, where significant resources are allocated to the theory of change development process, programme staff often get less invested in the theory of change as day-to-day demands of the programme start to take hold. It is not uncommon for programmes such as these to let drift their theory of change aspirations, only to review the theory of change on a yearly basis, as part of an annual review cycle. Reasons for this vary and often fall into the category of 'benign lack of use', as discussed by Perrin in this volume (Essay 19), including the use of a theory of change as a pro forma, rather than as a useful tool for programme implementation. In other cases, the theory of change gets abandoned as teams cannot see their activities reflected in the often-simplified programme theories of change, as suggested by Freer in this volume (Essay 16).

Supporting programmes' implementation teams to own the theory of change should be a key consideration when implementing a programme, so they use, reflect, and update their theory of change on a regular basis. This essay explores four practical ways in which programme leadership and external advisors can provide this support, which is born out of experience working with a range of programmes to improve how they learn from practice and adapt their implementation. Each of the proposed methods includes ways to integrate them into standard programme management processes, along with examples of programmes in which the methods have been trialled, and the benefits and challenges of doing so.

1. Develop theories of action for different programme teams

The first method that can be used to support uptake of the theory of change in programme implementation is the development of theories of action for specific parts of the programme. Funnell and Rogers (2011) define a theory of action as being what a programme expects to do to activate its theory of change. In practical terms, these are the expected pathways from the specifics of an intervention to the changes in behaviour of the groups that the intervention targets (Koleros et al., 2020; Mayne, 2017).

It has proven useful to develop one or more theories of action for each of the teams within a given programme, as programme teams are generally organised around groups of interventions reaching similar target populations. By providing a clear theory of action for each team, it is easier for the team to see their work accurately represented than in an overarching theory of change. Furthermore, a theory of action allows teams to see how their implementation work will fit into the wider change theory that the programme aims to unlock.

On a UK-funded social accountability programme in Pakistan, implemented by Palladium, the programme team developed theories of action for several parts of the programme. Using the Actor-Based Change Framework (Koleros et al., 2020), these theories of action included detailed descriptions of how each intervention was intended to lead to a change in practice among targeted actor groups. They also included key intervention assumptions: those conditions that would need to hold for the behaviour change to happen.

For example, the programme's policy and advocacy team worked predominantly with politicians and political activists to establish active local and regional advocacy platforms. This team developed specific theories of action that detailed the pathways through which their advocacy activities intended to change how political actors raised maternal and neonatal health issues to provincial governments. Meanwhile, the community outreach team worked mainly with local leaders and community organisers to bring the community together, and their theories of action centred around how their capacity-building activities intended to improve community organising and political engagement of citizens. As the

theories of action were specific to the work that each team conducted, the teams felt high levels of ownership of their theories of action and frequently updated them, as the programme learned from and adapted its implementation.

The overarching theory of change for the programme, in turn, described the ways in which the behaviour changes by reached actors would lead to wider changes in the system. The theories of action were distinctly linked to the theory of change, making it clear for each team to understand exactly what was expected of them and why.

One thing to bear in mind when developing specific theories of action for each team is to ensure that a team's performance is not equated to the validity of their theory of action. Where teams identify strongly with their theory of action, they may become so invested in proving it that they are blind to their own causal assumptions. It is important for programme leadership to foster a culture of reflective practice and learning, where teams can step back and observe to what extent the hypotheses in the theory of action held and openly admit where they did not, regardless of team efforts and without fear of failure.

2. Use theories of change to structure routine learning and adaptation

Another way to foster regular use of the theory of change is to explicitly include it in regular learning and adaptation reflection sessions. It is important to note, though, that not all components of a theory of change need to be considered at the same time and at the same intervals. For example, higher-level outcomes generally take more time to manifest, and they are unlikely to need amending frequently. On the other hand, programme activities are likely to need to change frequently and will need to be revisited regularly. To organise this, the theory of change hierarchy is a useful way to structure the learning cycles for each level.

The following is an illustrative example of what this could look like in a programme, and how it can link the theory of change hierarchy to the frequency of programme reflection sessions:

- On a weekly basis, programme teams discuss the delivery of programme activities. Attached to weekly staff meetings, the team discusses what is going well and what is not, making tweaks to the way these activities are planned and implemented.
- On a monthly basis, programme teams discuss intervention implementation experiences, in terms of the quantity and quality of the goods and or services (or outputs) the programme provides, whether these reached sufficient intended actors, and what the recipients' responses are to these (including uptake and satisfaction).
- On a quarterly basis, programme teams discuss changes in the determinants of behaviour, as well as actual behaviour changes (or outcomes) of the target actors of the interventions. The contribution of the intervention to these

changes is discussed, and evidence is collected and validated. Contextual changes and unintended consequences are also discussed quarterly.

- On an annual basis, programme teams discuss how any changes in the behaviour of the actors they worked with have affected the overarching change theory, and how the sum of the programme interventions has contributed to the intended impact.

Based on the nature and maturity of the programme, this schedule of sessions might be amended. For example, an adaptive programme may wish to more regularly reflect on the extent to which it contributes to actor-level behaviour changes, or a more mature programme might not need to discuss the delivery of goods and services monthly if these have already been honed over a longer period.

On the Global Monitoring, Evaluation and Learning (GMEL) Partnership for the UK cross-government Conflict, Stability and Security Fund, implemented by Integrity, this schedule was amended to fit the needs of the programme. The experimental nature of the programme meant that lots of smaller interventions (small bets) would be attempted simultaneously, which would make it impossible to discuss all interventions across the entire team. Rather, intervention teams have frequent—sometimes even daily—meetings on the delivery of the intervention and, at minimum, have quarterly reflection sessions that focus on the intervention theory of action. These meetings feed relevant information into a cycle of wider programme reflection, which includes quarterly reviews of groups of interventions, which discuss wider contributions to achieving the GMEL theory of change, emerging lessons from implementation, and proposed adaptations to the programme strategy and theory of change. Notes from these reflection sessions are then synthesised for an annual review and revision of the theory of change.

3. Link assumption monitoring to risk management

A third way of fostering active use of the theory of change on a more routine basis is to link assumption monitoring to risk management. At its core, a theory of change is a set of hypotheses (rationale assumptions) about how change will happen and what conditions need to hold true (causal link assumptions) for this change to occur. This means that, theoretically at least, risks and assumptions are closely related: an assumption not holding true over time is a risk to the programme. Yet, risks and assumptions are dealt with differently and by different teams within the same programme.

A good place to start is to incorporate all the assumptions that are identified in the theory of change development into a programme risk-mitigation structure or matrix. First, include all the causal assumptions (which could be referred to as programme logic risks). These appear at different levels of the theory of

change: at the lower end, there are risks associated with the programme inputs and activities not delivering the correct outputs (programme implementation risks). At the higher end of the theory of change, risks may include the identified outputs not resulting in the intended outcomes and impact (programme theory risks).

Next, it is important to consider the causal link assumptions that need to hold true for the causal logic to hold. On the Pakistan programme mentioned previously, one intervention was designed to use radio broadcasts to target young mothers with the intention of changing their health-seeking behaviours. A causal link assumption that was identified was that young mothers listened to the radio. However, a survey conducted by the team at the start of the programme suggested that only a small minority did. Had this not been identified, this would have been a major risk to the programme logic, yet this did not appear in the risk matrix at the start of the programme. Therefore, it is good practice to include these assumptions from the programme theory of change in the risk matrix to give a more complete picture of the development and delivery risks a programme faces.

Conversely, it is worth applying risk management principles to monitor these assumptions. In a risk matrix, the probability of a risk occurring is identified, as well as the level of the impact it would have if it did occur. Based on this, a risk-mitigation strategy is developed. Unfortunately, the same rigour is not usually applied in monitoring and managing assumptions. For example, as part of monitoring and evaluation services provided by Integrity to UK delivery in Somalia, we reviewed the theory of change of a programme in Somalia, which assumed that the political situation in the country would remain stable for the duration of the programme. In the volatile context of Somalia, this assumption was not likely to hold and, in fact, it did not. If this assumption would have been subject to the same level of rigour with which we think about risks, the assumption would have been refined or updated to take relevant political considerations into account.

4. Build the management information system around the theory of change

The final suggestion for incorporating the theory of change in programme management is through designing the programme management information system (MIS) in a way that it reflects the organisation of the theory of change. This way, it allows programmes to store and link information on programme management (including work plans and budgets), together with monitoring and evaluation information (including progress reports and documentation of results). Combining this will provide the programme team with much clearer information on what resources are dedicated to different parts of the programme and to what extent different parts of the programme are working or not.

We developed such an MIS on the Global Monitoring, Evaluation and Learning Partnership, referenced previously, and we organised it in such a way that it was closely linked to the programme's theory of change. We linked work plans and financial information to diarised delivery of activities, which are linked to specific intervention theories of action. This allowed us to see exactly when and how our services reached specific actor groups. It also facilitated linking evidence of changes in behaviour and changes in the system we operate in and linking these to interventions where we have evidence to support our contribution. Capturing this information together and ensuring that relevant linkages are captured opens a wealth of information and the uses for programme management and improvement are innumerable. We actively use the information collected in our regular reflections, and this has allowed us to conduct rigorous outcome harvests and real value for money assessments.

One potential downside that we have identified in this approach is that whilst making amendments to the MIS is not difficult, it does take time to implement fully. Given the adaptive nature of the programme, this has at times proved quite burdensome. However, overall, we feel the benefits as described have far outweighed this cost of upkeep.

Conclusion

In this essay, we present four practical ways in which the theory of change goes from a shelved document to a central part of programme management:

1. Develop theories of action for different programme teams.
2. Use theories of change to structure routine learning and adaptation.
3. Link assumption monitoring to risk management.
4. Build the management information system around the theory of change.

The success of these approaches in large part is due to their continuous application and relies heavily on the capacity of the programme to support these processes. For evaluators, this might mean an engagement strategy with the programme that is based on capacity building, facilitation, and advisory support, rather than a fully external role in which engagement peaks at the start and end of the programme. In all the programmes discussed (monitoring and) evaluation advisors were either fully embedded in the team or provided routine support; this has resulted in the theory of change living up to its potential for these programmes.

References

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory*. John Wiley & Sons.

Koleros, A., Mulkerne, S., Oldenbeuving, M. L., & Stein, D. (2020). The actor-based change framework: A pragmatic approach to developing program theory for interventions in complex systems. *American Journal of Evaluation*, *41*(1), 34–53.

Mayne, J. (2017). Theory of change analysis: Building robust theories of change. *Canadian Journal of Program Evaluation*, *32*(2), 155–173.

Stein, D., & Valters, C. (2012). *Understanding theory of change in international development*. London School of Economics. http://eprints.lse.ac.uk/56359/1/JSRP _Paper1_Understanding_theory_of_change_in_international_development_Stein _Valters_2012.pdf

Valters, C. (2015). *Theories of change: Time for a radical learning agenda in development*. Overseas Development Institute. Retrieved from https://odi.org/en/publications/ theories-of-change-time-for-a-radical-approach-to-learning-in-development/

Vogel, I. (2012). *Review of the use of 'theory of change' in international development*. Department for International Development.

19 Theories of change: Who needs them?

Or: What evaluators can learn from opera

Burt Perrin

Introduction—A universal need?

Theories of change represent a potentially useful technique. But are they really needed—by all programmes, all evaluations, in all possible situations—as they increasingly are being promoted?

An operatic story, Donizetti's L'Elisir d'Amore (The Elixir of Love), may present some useful lessons for theories of change in the worlds of evaluation and programme planning. In this opera, Nemorino is in love with Adina. He is a nice guy, but penniless, simple with limited skills; he really has little to offer her. In the context of evaluation, one might think of a programme that is well intentioned but with inadequate resources and limited competencies—yet with lofty, seemingly unrealistic objectives.

A military battalion arrives in the village—and Adina announces that she will marry the dashing officer—later that day! (Our programme is going to be shut down imminently!). Nemorino, desperate, turns to the charlatan Doctor Dulcamara, who also just arrived in town, pedalling a secret potion that is good for everything— indeed the way all programmes often are exhorted to develop a theory of change. He tells Nemorino to give him all his money and to drink the bottle.

And—it works! Suddenly, Adina realises that she loves Nemorino and agrees to marry him! To be sure, there was what social scientists would call a confound: word just reached the village that Nemorino has come into an unexpected large inheritance!

In reality, the potion was red wine, which in the right circumstances and the right quantity can certainly be good for many things. Even, perhaps, for affairs of the heart, although one would think generally when shared with the object of one's desire rather than quaffed down all by oneself (a participatory approach, I will come back to this later). But clearly, wine is not the answer to everything; it is best consumed under certain circumstances, and it sometimes can make things worse.

Implications for theories of change

So, what does this mean for theories of change? Programmes and evaluations— regardless of status, size, shape, or situation—are increasingly exhorted or

indeed required as a condition of funding to develop theory of change models, often in a very prescribed format. Does this indeed make any more sense than to proclaim red wine as the solution to everything?

I appreciate that statements such as the above, as well as others that follow within this essay, may seem critical of theories of change. To be clear and to put this into context, I have long been an advocate of theories of change, frequently using these myself and recommending them to others. And I still do. But to say that every programme must have a theory of change, in a designated format, makes as much sense as to believe that Dr Dulcamara's secret potion, or red wine, is good for everything. Too often, this is how it is portrayed.

Accordingly, I feel that a critical consideration of theories of change is needed, as this book is trying to do, in essence to take a realist approach (for example, Pawson & Tilley, 1997) to theories of change to identify the circumstances under which a given theory of change approach might be appropriate or not.

The reality of theories of change in use

I have seen many different theory of change models and approaches, arising in part from my experience as an expert advisor and quality assurance reviewer of numerous evaluation plans and reports, on behalf of a wide variety of governments and international agencies. There have indeed been some good theory of change approaches that have been useful and have added value. Unfortunately, however, the vast majority have not. As Sanjeev Sridharan (2021) observed, there often is a gap between the theory and the practice of theory of change. Models that have not been useful fall into two subcategories.

Benign lack of use. The first category, which applies to most of the theory of change models that I have seen, is benign. This means that while they have not added value or have been used, they have not done any harm, other than representing a waste of time and resources, and perhaps contributing to cynicism.

There are two common ways in which this can happen. The first refers to a theory of change model that appears in a report, possibly referred to once or twice, but it has not played any meaningful role. It is essentially pro forma.

The second subcategory of benign models concerns those that hardly anyone understands or can relate to. Perhaps the model is overly complex and convoluted. As the saying goes, a picture can be worth a thousand words. But, if a theory of change model requires thousands of words for anyone to understand it, then how helpful is it really?

What makes for a useful model? It can be tricky to get the balance right between an overly complex and incomprehensive model versus one that is overly simplistic and that might misrepresent what a programme is trying to do and its context. As Einstein has said: 'Everything should be made as simple as possible, but no simpler'.

Stakeholders are less likely to understand theory of change models when they have not been involved in its development, such as the all-too-common situation where a theory of change model has been prepared mainly or entirely by an evaluator, often in isolation. Worse, this may result in a model that represents how the evaluator views the programme and not how those involved in it expect it to work.

Negative impacts of theories of change. The previous section depicts situations where theories of change may not especially add value, but at least they do no harm. In contrast, there are other situations where theories of change can be negative or even harmful. Unfortunately, this may be becoming more common. This should be no surprise. For example, while evaluators often assume that evaluation is always good, other books in this series (Furubo & Stame, 2018; Perrin & Tyrrell, 2021) have documented that the value of evaluation cannot be assumed.

There are various ways in which theories of change might have a detrimental effect. The first, and potentially the most pernicious, is when a theory of change becomes a rigid input-output model, with indicators and targets identified in advance. This type of model is, at best, a glorification of the logframe: It represents a bastardisation of the theory of change concept, which—as Freer discusses in his essay—is intended to articulate assumptions about how a programme is expected to work, including mechanisms, interactions with other interventions and intervening factors, and potential feedback loops. Indeed, this represents the same phenomenon that has afflicted the logic model that initially was much the same as what is now called a theory of change but, over time, was to a large extent reduced to a confined input-output model with little room for emergence or explanation.

A related problem can arise when a theory of change model makes no room for uncertainty. In contrast, a meaningful model should be dynamic rather than static, adapting in response to unexpected events that in a rapidly changing context might not have been identified, or even knowable, in advance (with COVID-19 representing one example). A static model runs the risk of locking a programme—and an evaluation—to previously conceived expectations about what might occur; this may turn out to be wrong or out of date and can lead to tunnel vision and to misrepresentation of current reality.

Some theory of change models are aspirational in nature, representing statements of hope rather than reflecting underlying assumptions or reality. This can lead to distorted programme priorities and, potentially, to an evaluation focusing on inappropriate questions. It can also make adaptations to the model difficult if it becomes a quasi-policy document.

Table 19.1 summarises some key considerations that I very briefly discuss below.

1. **What is the theory of change for the theory of change?** In what ways could the development of a theory of change model be useful? I suggest that this be

Table 19.1 Some key considerations when contemplating the use of a theory of change

1. What is the theory of change for the theory of change?
2. Start with the desired impacts rather than with the intervention.
3. Engage stakeholders in articulating a theory of change.
4. Consider narratives, stories, and diagrams.
5. Build in provision for unexpected factors and effects.
6. Use a bespoke model appropriate for the particular intervention.

considered before jumping headfirst into work on a theory of change. If the expected benefits and values of a theory of change are not clear, then there is little point to proceed. Similarly, there would be little point to undertake a theory of change approach if it is not likely to be used in practice.

Failure to think through in advance the potential benefits and use of a theory of change model may be responsible for the large percentage of theories of change that are, at best, pro forma. A theory of change approach represents a technique, an output, rather than an end. It is imperative that it is viewed in this way, otherwise the end result might be a beautiful diagram that no one uses.

2. **Start with the desired impacts rather than with the intervention.** Theory of change models, ideally, should show the expected link between what is being done and its impacts. Theory of change models starting with the intervention may lead to narrowness, failing to consider the inevitable interactions with other factors and limiting the ability of a programme to have a real influence on what is most important.

   The best way to avoid this pitfall is to start at the impact level and then work backwards, identifying multiple factors—including but not limited to the intervention—that might be contributing to impact.

3. **Engage stakeholders in articulating a theory of change.** As suggested previously, it invariably is best to engage stakeholders in the development of the theory of change to produce a shared perspective and understanding of what the model is saying and the assumptions of those engaged in the programme regarding underlying assumptions and hypotheses. Otherwise, there may be little resemblance between what an outsider thinks the programme is (or should be) and what it actually is trying to accomplish.

   In my experience, the process of stakeholders collectively thinking through what their intervention is trying to achieve and in what ways can be very powerful, even more important than the resulting model. This represents an example of process use of evaluation (for example, Patton, 1997).

4. **Consider narratives, stories, and diagrams.** There often is an assumption, indeed one I have made myself at times, that a theory of change must be portrayed through a diagram. But not necessarily. I now realise that there is strong potential for narrative descriptions that in some cases can be deceptively simple—while still making provision for uncertainty and emergence. Indeed, even with more complex models, an accompanying narrative might add to understanding.

Another narrative approach with the potential to give a flavour of how a programme is expected to work is stories, for example, *Evaluations that Make a Difference* (Perrin et al., 2015). A stories approach potentially could also provide greater opportunities for engagement of stakeholders. Indeed, John Mayne (2015) has observed that a good theory of change makes for a good story.

5. **Build in provision for unexpected factors and effects.** As indicated previously, virtually all public-sector initiatives operate in a context of uncertainty. Some theory of change models assume that all possible outcomes, intervening variables, and pathways can be fully knowable in advance; this runs the risk of becoming increasingly irrelevant.

If using a diagram, one might insert links or boxes at various places to indicate possible unexpected influencing factors and effects, and where interactions with other interventions or factors might, perhaps, be anticipated. In any case, models should be viewed as dynamic rather than static, and they should be reviewed and updated on a regular basis.

6. **Use a bespoke model appropriate for the particular intervention.** This should be a no-brainer: When developing a theory of change model, it should reflect the situation and expected impact pathway of the intervention in question, which varies from situation to situation. As Funnell and Rogers (2011) illustrate, a wide range of potential models is possible. When stakeholders play an active role in the development of a theory of change, a model that makes sense to them is most likely to be a useful tool, in contrast to standardised formats that are frequently mandated or imposed.

When to consider using a theory of change

As I have suggested, theories of change potentially can represent a valuable aid to programme planning and implementation and to focusing evaluation. But this is not automatic, and theories of change, as indeed some evaluation undertakings, have the potential to do harm.

The first, and arguably the most important, criterion when considering taking a theory of change approach is to identify how it might be used and by whom. Use need not be dramatic. It may, for example, aid in enhancing evaluative

thinking and in raising questions about how a programme is expected to work. Following are some common ways in which a theory of change might be worth considering. This, to be sure, does not represent an exhaustive list.

• When it can help articulate hidden assumptions about how a programme is expected to 'work'—and when this information would be useful. As a corollary, a theory of change approach can be valuable if it can help build a common understanding of the programme's underlying assumptions. If this is not possible, it can identify competing theories about how a programme might work, which can then be tested through evaluation.
• When it helps to suggest which possible evaluation questions might be appropriate to explore at given points in time. A common mistake is to evaluate the wrong questions prematurely.
• When it demonstrates interactions and possible feedback loops, both internally and with external factors and other interventions. This can also be a means of identifying how a given intervention might be contributing towards impact down the road, which rarely happens without interaction with others.
• When it is treated as a dynamic rather than a static tool that can change over time and acknowledges the reality of emergence.
• When it is understandable—and considered relevant and helpful—by key stakeholders. If it makes no sense to stakeholders or if it does not represent the reality of the programme as they know it, positive use would be unlikely, at best.

Conclusion

With the right approach and in the right circumstances, a theory of change approach has the potential to serve as a powerful and useful tool. For this reason, I have long used and promoted this technique—but it clearly is not needed in all situations. Unfortunately, most theory of change models that I have seen either have not been used at all or they may have been detrimental. When this approach is used, it is important that it be done right, such as adhering to the considerations and guidelines identified above. There is little point in proceeding with a theory of change approach if it is not likely to be used. Its benefits are not guaranteed, and it can even do harm. Theories of change are not a magic potion, such as in l'Elisir d'Amore. At best, it represents a technique, an output, rather than an end. Like a good bottle of red wine, it often can be very desirable—but it is not always appropriate, helpful, or needed.

References

Perrin, B., Zorzi, R., Rodriguez-Bilella, P., Bayley, S., Yakeu, S., & De Silva, s. (Eds.) (2015). Evaluations that make a difference. Retrieved from: https://evaluationstories. wordpress.com/evaluation-story-publications/

Funnell, S. C., & Rogers, P. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. Jossey-Bass.

Furubo, J.-E., & Stame, N. (2018). *The evaluation enterprise: A critical view*. Routledge.

Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, *30*, 119–142.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Sage Publications.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage Publications.

Perrin, B., & Tyrrell, T. (2021). *Changing bureaucracies: Adapting to uncertainty, and how evaluation can help*. Routledge.

Sridharan, S. (2021). *Moderator's comments during a webinar on theory of change*. The Evaluation Centre for Complex Health Interventions.

20 How to ensure no-one uses your theory of change

Lessons from the front lines of theory of change facilitation and possibilities for renewal

Mary Tangelder

The origin story of the theory of change is that leaders at the Aspen Roundtable on Community Change pioneered it in the 1990s as a design tool for tackling complex challenges, such as systemic racism or promoting equitable economic development (Corner, 2019). But I have always imagined that the theory of change emerged from a deep place of yearning to imagine new worlds, just like six-year-old Simon in the beloved 1970s BBC show Simon in the Land of Chalk Drawings. On his way to school, Simon idly draws pictures on the neighborhood fence, only to quickly discover his power to bring his chalk drawings to life. In each episode thereafter, Simon continues to create new worlds and, despite his best intentions, regularly triggers unintentional consequences, such as a rampaging dinosaur or a measles epidemic.

Like Simon's chalk drawings, the theory of change tool was purposely designed to spark imagination about different, better worlds and spotlight what happens when we intervene. Like a map, a theory of change is an explicit journey from here to there, inviting viewers to locate themselves within that journey and start to make strategic choices about where to go next. The design process invites us collectively to surface our assumptions, highlight blank spaces in the logic, and ultimately influence the journey we choose to take (Aragon, 2010).

A growing number of funders, particularly in the private philanthropy space, are adopting systems change thinking and methodology to tackle the root causes of large-scale problems. To support this shift, theories of change are increasingly used as a design tool to shift the focus from short-term, discrete activities to highlighting possibilities for collective contribution to longer-term outcomes and, ultimately, transformational change.

Or so the story goes

Despite the theory of change's oft-quoted benefits, my experience has taught me that there often is no better way to vacuum-suck energy from a room than for someone to float the dreaded phrase: 'I think we need a theory of change'. And rarely do programme teams use it to guide learning conversations, pin it on

their wall and refer to throughout programme implementation to learn or adapt or revisit their theory of change before it is resurrected by an external evaluator.

Why do theories of change have such a tarnished reputation among those at the front lines designing, commissioning, and implementing social change programmes? In other words, among anyone who is not a programme evaluator?

This essay draws from my own experiences facilitating or observing numerous theory of change exercises for programme design and evaluations with a range of organizations globally in the non-profit sector, including NGOs as well as government and multilateral organizations. It examines points of resistance and provides possibilities for leveraging and deepening the application of this tool for tackling persistent, large-scale, seemly intractable problems.

Roots of Resistance

One morning, two small fish taking a leisurely swim are passed by an elderly uncle. 'How's the water today, boys?' he calls out cheerily. The boys reply, 'Fantastic!' After the uncle passes, one fish turns to the other and asks: 'What's water?'

Whether we admit it or not, we are all swimming in theories of why things are the way they are, how change occurs, and what should be done about it. The implicit, often unspoken theory of change that has dominated decision making among funders goes something like this: fund good people, find and invest in good ideas. Oh, and be innovative! Better yet, sustainable!

When I have asked participants in a theory of change exercise why the term *theory of change* triggers a near-choreographed slump, I have been told that the initial resistance stems from the uneasy relationship we have with the word *theory*. For some, *theory* has colonial overtones, evoking a picture of inaction governed by remote evaluators and academics who spin theories but never have to live in the world of hard choices. Those addressing complex social challenge understandably do not want to live in the *Land of Theory*; they want to live in the *Land of Action*. Why take precious time to talk theory when people's lives are at risk? Who has time, who is this really for other than a proposal requirement, and what is the point?

For many people at the front line of social change efforts, their first encounter with a theory of change might be at the proposal gate, where a funder may require a detailed theory of change. Best intentions aside, this request may be disempowering for partners who have not been versed in development-ese and are unfamiliar with industry design tools. Similarly, having an external evaluator arrive at a midterm evaluation to flag all the potholes in a theory of change can feel exceptionally disempowering to programme staff, who may have course-corrected and renovated their theory of change along the way, but their adaptations and new insights simply have not been documented. People who are commissioning, implementing, or designing social change programmes are

often working under a deep sense of urgency, and it is difficult to prioritize documenting all the learning and adaptations when it feels like the house is on fire.

To make matters worse, the theory of change tool itself awkwardly straddles visual and narrative worlds. Visually, they are often a bewildering mix of shapes and lines that makes little sense except to the person who created it. To shift the notion that change is a linear process of cause and effect, theory of change advocates will often layer an increasing number of connections, arrows, and diminishing font size; the result can be crippling anxiety about how to make strategic choices.

At the same time, an uncomfortable feature is that a theory of change attempts to tell a story about change without showing us any people (and clip art does not count). Visions of new and better worlds are reduced to mechanical boxes and arrows that emphasize inputs and outputs rather than people and relationships, a visual rendering of change that is miles away from the animated and dynamic worlds Simon experiences. We simply do not experience life this way.

A potential way forward

In my experience, a theory of change exercise has been a powerful way to bring clarity around a multifaceted, complex challenge and build pathways to a shared vision of change. It has also been a tool for 'learning in action', illuminating what assumptions to test and what data to prioritize to help us learn, adapt, and accelerate. Yet it is evident that the theory of change is in serious need of disciplined application and rebranding.

Start by involving the experts in the analysis of the challenge

Although boxes and arrows might suggest otherwise, a theory of change is not intended to engineer a simple solution, but rather to guide purposeful action and learning within complexity. Learning through complexity requires first stepping away from the jargon and heavy-handed PowerPoint slides and taking time to collectively understand the nature of the challenge. A critical aspect of this step requires involving experts in the analysis—those who have deep lived experience navigating the cracks of a broken system: parents navigating health, education, and child-care systems; farmers working in food systems; and so on.

Spend time on articulating the vision and goal, even though it hurts

A theory of change brings a goal statement sharply into focus—a step that seems simple at first but is notoriously tricky and much more time-consuming than anticipated. Annie Duke, a former professional poker player and cognitive behavioral scientist, uses an example from mountaineering to explain why getting the goal statement right is so critical before embarking on a journey, and why success depends on making brutal choices in the face of uncertainty and limited

data. A mountaineer climbing Mt. Everest typically cites their goal as 'Get to the top'. However, mountaineers who define their goal as 'Get to the top *and* come back alive' are much more likely to make better choices (Shankar, 2021).

Duke also argues that the goal statements also fundamentally shape our narrative of success and whom we elevate as heroes. For example, if we collectively understood the goal as 'Come back alive', we might start to tell a different version of the hero story—not simply elevate the lone hero who completed the journey at great cost to self and others. Rather, we might celebrate leaders who use the best evidence at hand to make difficult, unpopular decisions and change course, such as a leader's decision to switch back despite close proximity to the top (Shankar, 2021).

Incentivize and invest in learning and adaptation

While a simple, predictable challenge (putting out a fire) requires a simple plan and can often be fixed with short-term immediate solutions (proximity to a fire hose), addressing a complex, unpredictable, and evolving challenge (farming or raising a child) requires different design tools that can take years to address. As every farmer or parent knows, it involves rigorous attention to moving targets and shifting contexts, comfort with uncertainty, humility to engage in more learning that you have ever done before, confidence to abandon plans and change course, and capacity to mobilize support and build relationships across multiple systems. And there simply is no finish line (Glouberman & Zimmerman, 2002; Ricigliano, 2022).

When addressing complex versus simple challenges, a theory of change can reinforce this notion that learning and adaptation is really the only choice. Yet funder reporting requirements and mechanisms that hold grantees accountable for sticking to the plan can significantly stifle efforts to navigate complex change efforts through continuous learning and adaptation. Short one- to three-year funding cycles further incentivize programming oriented to address symptoms versus systemic causes, and social change organizations are left lurching between heavy reporting requirements and chasing the next grant. Funders can turn this around by expanding timeframes, providing resources for reflective practice, and shifting notions of accountability by asking how they are shifting and adapting in response to what they are learning—rather than assessing fidelity to their initial plan.

Incentivize and invest in learning, for real

Like watching someone complete a jigsaw puzzle, it is excruciating to observe a theory of change design process from the sidelines, yet this is often people's first experience with it. The only way to accelerate the design of the theory of change is to go it alone, yet a theory of change is a deeply participatory, co-creative exercise, which simply takes time. More than time, it requires facilitation expertise to get people to roll up their sleeves; engage in a messy, unscripted process; negotiate competing perspectives; allow space for diverse viewpoints and

innovative thinking to emerge; and stimulate courageous conversations about what can be done, by whom, for whom.

Ultimately, what this tool is called and what the output looks like simply does not matter. What matters is intentional design and commitment to deep learning that will enable us to make good choices throughout the journey. To start, a funder can ask applicants to explain, in their own words in dialogue or written form, how their activities will get us from here to there. Funders can then co-design a more detailed theory of change with partners post-contract and bring in skilled facilitators to guide the design process. A facilitator can also attend to the risk of drowning in a pool of analysis by using the GEPO principle: good enough, push on, and guide sense making and collective learning throughout the programme cycle: What are we learning about the political, social, and environmental landscape we operate in? What are we learning about the ecosystem of actors? Who else is working in this space and what is our unique contribution? What should we adapt or where do we shift course?

Create multiple views for the right audiences

Just like the painful exercise of watching someone do a jigsaw puzzle, framing a puzzle or distributing a detailed theory of change for public consumption seems completely beside the point. Instead, multiple views can be created for different audiences who have different communication needs, such as a simplified, summarized version for internal leaders (for example, board, executives), a detailed version for programme managers to help navigate decision making, a compelling visual to stimulate dialogue and mobilize collaboration, and so on.

Consider engaging storytellers and artists

Boxes and arrows can only go so far in telling the complex story of a challenge and the proposed solutions. But art can take us farther, bringing emotional resonance and drawing us more deeply into a narrative of change, particularly when participants are leading the creative process. For example, Slum Dwellers International engaged a group of young people in a two-day theory of change workshop in Nairobi and encouraged young people to communicate their theory of change in a way that made the most sense to them. The result was a compelling role play and vivid artwork that showcased a theory of change told by young people, for young people, and for their community (Impact Lab, 2021). Increasingly, more attention has been given to the oldest tool we have—storytelling—as a way to cast light on fault lines and fractures within a system and highlight disruptors, outliers, and possibilities for new futures (Saltmarche, 2018).

Calibrate the level of effort to the purpose and moment

Not everyone or every moment needs a theory of change at the same level of detail, nor does every process need to be facilitated in the same way. Designing

a new strategy, launching a new initiative, and undertaking an evaluation are all moments when an in-depth theory of change process may be required (Rogers, 2008). At the same time, a pared-down process could be used as part of a sense-making and learning activity to bring teams together at pivotal moments in a team or organization's journey.

A final note

Once when speaking to a group of kindergarten children about indicators of quality education, I asked how I would know if a school was a good school. Turns out they knew exactly what to look for: 'Watch and look at how fast children run into school at the beginning of the day and back in when recess is over. And how quickly kids line up for things'. I then asked who else I should speak to, and they shared: 'The janitors. They watch and see everything and know about everything that happens'. It was yet another reminder of how often we underestimate the ability of people on the margins to participate in analyzing a challenge and shaping solutions and measuring success.

Tactics and facilitation notes aside, whole-hearted listening is still our most powerful tool for charting a way forward.

References

Aragon, A. (2010). A 'systemic theories of change' approach for purposeful capacity development. *IDS Bulletin*, *41*(3), 87–99. www.researchgate.net/publication /249500647_A_'Systemic_Theories_of_Change'_Approach_for_Purposeful_Capac ity_Development

Corner, J. (2019). Systems change and philanthropy. *Alliance Magazine*, March. www .alliancemagazine.org/feature/systems-change-and-philanthropy/

Glouberman, S., & Zimmerman, B. (2002). *Complicated and complex systems: What would successful reform of Medicare look like?* Discussion Paper No. 8. Commission on the Future of Health Care in Canada.

Impact Lab. (2021). *Youth from Nairobi's informal settlements generate a theory of transformation* [Workshop Report]. Know Your City TV, November 29 – December 1.

Ricigliano, R. (2022). The complexity spectrum: Systems change is not for everyone, but making a good choice is. *Too Deep*, January 5. https://blog.kumu.io/the-complexity -spectrum-e12efae133b0

Rogers, P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29–48. https://journals.sagepub.com/doi/pdf/10 .1177/1356389007084674

Saltmarche, E. (2018). Using story to change systems. *Stanford Social Innovation Review*, February 20. https://ssir.org/articles/entry/using_story_to_change_systems

Shankar, M. (Host). (2021) Annie Duke: The science of quitting. *A Slight Change of Plans*, November 1. www.annieduke.com/podcast-a-slight-change-of-plans/

21 Why do we have theories of change of the programme intervention but not of the intervention that is the evaluation?

Bob Williams

The consequences of evaluations

There is a point in my systemic evaluation design workshops that I approach with foreboding. It never goes well and I have never been able to work out how to make it work well. This is possibly because the point exposes an embarrassing weakness and contradiction in many evaluators' practice. The systemic evaluation design workshop is case-based, and at this point of the proceedings, I am prodding participants to identify the *consequences* of an evaluation. What I invariably get is puzzlement and confusion. The room is suddenly against me.

The history of discussions around evaluation consequences is intriguing. There is an acceptance going back many decades that evaluation is a purposeful intervention (for example, accountability, improvement, and learning). In other words, it is an action that has consequences. Those consequences may be intended or unintended—and almost invariably a mixture of the two. There is also a vast literature on how and why evaluations can and ought to be used to achieve these purposes. The discourse around *use* is ubiquitous, and the idea of evaluation usefulness is virtually beyond critique. It would be a career-limiting move for any evaluator these days to state they were not utilisation focused. On the other hand, evaluation use is only part of the consequence story. Use is only an action. Usefulness depends on your perspective, but consequences can hurt or heal.

Specific discussion on the *consequences* of evaluation has been relatively muted. There was a brief period when it looked as if that may change. In the early parts of this century, Mel Mark, Gary Henry, and George Julnes wrote extensively about evaluation consequences. It was the topic of Mel Mark's 2006 American Evaluation Association Presidential Address. Mel Mark's ideas are summarised in Table 21.1.

In recent years the discussion has drifted away from consequences in general and morphed into positional or power concerns (for example, indigenous issues and gender) and more methodological ones (for example, Blue Marble Evaluation, Developmental Evaluation, Outcome Mapping, and Outcome Harvesting).

Summing up—any jobbing evaluator would have to have been down a very deep rabbit hole for many years to avoid coming across debates around

Table 21.1 Summary of evaluation consequences from 2006 American Evaluation Association Presidential Address

Consequence	Personal	Interpersonal	Institutional
Evaluation influences behaviour	Affects the way in which individuals work	Affects the way in which people work together	Affects the way in which the institution works
Evaluation affects 'thinking'	Affects the way in which individuals think	Affects attitudes towards working with each other, or what other people do	Affects how the institution values certain kinds of ways of thinking. Changes in general strategy
Evaluation affects organisational structures and processes	Enables people to work more effectively or efficiently	Enables people to be able to work together more	Enables ways of allowing people to work better, individually or collectively
Evaluation affects broader goals, motivations, and aspirations	Encourages people to feel good about what they do or promote improvement	Enables people to be respectful and understand what each other do	Affects formal goals, values, and aspirations
Evaluation has consequences beyond the immediate interests of the project	Encourages people in other agencies to adapt and adopt the ideas and methods that underpin the project	Promotes the benefits of collaboration and cooperation in a more general sense	Inspires organisations to follow (or avoid) similar ideas

evaluation consequences and specific values-driven evaluation approaches with an emphasis on social and cultural justice. Evaluators really have no excuse to be unaware of the ways of addressing the consequences of their evaluative interventions. So, it intrigued me why I was faced with blank stares and not infrequent pushback when I suggested a discussion around identifying desirable or undesirable evaluation consequences.

Evaluation use, evaluation consequences, and theories of change

But a quick step back. At this point, you may be puzzled about where exactly theories of change fit into this picture.

A few years ago, during a discussion on the application of principle focused evaluation at an American Evaluation Association conference, it seemed that for evaluators under the age of 30, the production of a theory of change was a

mandatory part of every evaluation. When those of us with greyer or no hair pointed out that we had done perfectly good evaluations without using theories of change, it was as if we had proposed making a margarita and skipping the tequila. Theories of change are these days (rightly or wrongly) considered to be part of the very fabric of evaluation. As Burt Perrin also notes in his essay in this volume, it is virtually impossible to have any discussion about evaluation design without discussing theories of change.

Clearly, identifying and designing for the consequences of any particular evaluation depends on some kind of implicit theory of change (small capitals). But why, in contrast with the theory of the intervention being elevated into a full-blown theory of change, does the theory of change embedded in the evaluation remain implicit, unexplored, and unchallenged? If theories of change are considered virtually compulsory in evaluation these days, why is it not standard practice to undertake a theory of change for the evaluation itself? We force those involved with the intervention to expose their theories of change and leaves them open to critique and assessment. So why not ours?

Returning to the topic of my workshop. When I suggest that perhaps we should construct a theory of change to explore the potential consequences of an evaluation, the temperature of the room definitely rises. Participants tend to provide all kinds of reasons why evaluations are not like 'programs', and thus it is not appropriate. Or they argue that the consequences of evaluation are far too unpredictable. Or the context is too changeable, too political, or too unknown. I try to point out that if that is the case for the evaluation, then in all likelihood it will be even more so for the programme. At which point I generally back off—my workshops are fundamentally not theory of change workshops and the topic becomes a distraction. But the question at the start of this essay remains hanging in the air:

Why do we have theories of change of the programme intervention but not of the intervention that is the evaluation?[1]

For many years now, I have used a data analysis approach developed from Activity Systems that identifies and then seeks to explain (rather than explain away) contradictions.[2] The attractive thing about contradictions is that both sides of a contradiction are 'true'. Neither contradiction can be dismissed easily, so the rationale for the contradiction requires explaining—not explaining away. Explaining how both sides can exist leads to a deeper understanding of a situation. So, when I am faced with a puzzle, I immediately seek to identify any

---

1   As an exception to this generalisation, see Hansen et al. (2013). On the other hand, a recent otherwise excellent paper on constructing an evaluation theory of change (Bitar, 2022) makes no mention of evaluation impact, results, or consequences in its description of 38 components of a theory of evaluation.

2   For more details on this approach, see Williams (n.d.).

underlying contradictions and then see if I can understand what might be going on behind those contradictions.

The above discussion exposes a simple contradiction:

On the one hand, theories of change are considered essential to the evaluation of interventions.

On the other hand, evaluation interventions almost never have their own theory of change.

I want to be very clear here. I am focusing on an apparent contradiction that I have observed about a particular use of theories of change. I am not implying that evaluators are uninterested in the usefulness of their evaluations, but that theories of change appear to be a rarely used to tool explore that use.

So… why the apparent contradiction?

What follows is pure speculation based on no evidence other than personal observation and experience. It is not intended to be comprehensive either. Consider it an agenda for debate.

First let me first try to see if I can explain the contraction away. My response is in *italics*.

The use of theories of change for evaluations is commonplace. My experience is an aberration. My observations have no validity and thus there is no contradiction. At most, my observations are exceptions to a generalisation. *Basically, analysis of contradictions is interested in qualitative rather than quantitative analysis. Even if 99% of the data says one thing, the other 1% still needs explaining because something is going on that might be worth exploring.*

So now let's seek to explain the contradiction. Again my responses are in *italics*.

1. Based on their experience or worldview, the evaluators intuitively know what the consequences will be. So they see no need to expose their implicit theory of change to the world. *But isn't exposing the logic of impact to broader discussion one of the reasons that evaluators argue for the use of theories of change? Isn't it a means of checking mental models and avoiding confirmation bias?*
2. The evaluators have a sophisticated perhaps emergent approach to the consequences of their evaluation. However, because the approach cannot be reduced to four sets of boxes and arrows, the evaluators do not consider it a theory of change, so they do not make it public. Patricia Roger (Essay 10) touches on the 'Four Box' issue. *That is no particular fault of an individual evaluator but an indication of the evaluation field's narrow understanding of what theories of change comprise and how to express them. We all bear responsibility for this state of affairs and, hopefully, this volume will expand our understanding. But many of the training programmes I have seen are*

depressingly narrow and rigid on what constitutes a theory of change and how to express it. That means evaluation has a capacity development issue to confront.

3. The evaluators consider 'consequences' as adequately covered within discussions around intended use by intended users. *True but 'use' is an action— not a result and certainly not an explicit theory of change.*

4. The evaluators do not consider their evaluation as an intervention. Since theories of change relate to interventions, constructing one for the evaluation is inappropriate. *Evaluations are always interventions; they change the status quo even a little bit.*

5. The evaluators sweat blood producing a theory of change for the intervention and just cannot face the prospect of constructing one for their evaluation. They feel they intuitively know how to make the evaluation useful anyway. *See #2.*

6. The evaluators do not really believe in theories of change and only construct one because it is expected or insisted upon by the evaluation commissioner. This is discussed in considerable detail in Gordon Freer's essay in this volume (Essay 16). *We have to search our sense of professionalism and pragmatism on this explanation. Should we be doing things we do not believe in just because our paymaster does? This is part of a much wider debate around evaluation practice and our professionalism.*

7. The evaluators see themselves primarily as having no control or authority over the way in which their evaluation is used or the consequences of their evaluation. So what is the point of constructing a theory of change? *That substantially depends on what the evaluators understand as a theory of change and its purpose. If it is a theory of change that is rigidly mechanistic and the purpose is primarily ontological, then the evaluators' judgement is probably correct. If they understand theories of change as essentially being mental models—a bunch of assumptions and beliefs that benefit from being exposed and discussed—then there is no excuse. As this publication displays, the range of theories of change is much richer than that. For instance, there is a whole world of system modelling out there to be drawn on and much more. Again, evaluation has a capacity development issue on its hands.*

8. There has been passive and aggressive behaviour from evaluation commissioners. Commissioners often have very good reasons for not declaring their hand about the potential consequences of the evaluation. Organisations are deeply political institutions. Commissioners are going to state openly that one of the consequences of an evaluation is that they can cut a program's budget by half, or the consequences are not going to be that the entire human resources division is going to be reorganised. And even if they are happy to tell you in private, they are not going to be happy with the evaluator announcing that to the world. *I am willing to give the evaluator some slack on this. I have encountered this issue on many occasions. Whether he, she, or I should*

have accepted the commission in the first place is a professional issue, but once an evaluator has sipped from the Devils cup there is no going back. Someone once advised me that the key is to be aware of the politics of a job but not get involved with them. I have not always taken that advice and generally lived to regret it.

9. There is a trend by commissioners to woefully underfund evaluations relative to their ambition (for example, expensive and time-consuming social research standards and requirements on crazy budgets and timescale). Evaluators just do not have the time or the money to devote to often detailed negotiations and discussions around evaluation consequences. *I am not prepared to give evaluators any slack on this. It is a question of professionalism—if you cannot do the job adequately, then you should not have signed the contract. If you hope that you can sneak things in later, possibly by inserting a 'scoping phase', then you are being hopelessly naïve.*

10. The evaluation commissioners do not have a clue about what they want from the evaluation. Even if they have an idea of how the evaluation might be used, they may have no idea how useful that will be or what consequences they desire out of that use. *We have all been there many times, but surely the first job of an evaluator would be to guide stakeholders through that process.*

And what about me?

All of which begs the question, what do I do? It is a combination of reasons #2, #7, #8, and #10. And, to be utterly honest, I have been guilty of #9, as well.

Am I committed to exploring the consequences of evaluations? Clearly yes. I am not a huge fan of the literal interpretation of 'intended use for intended users' (much too deterministic in complex environments). Instead, I generally ask key stakeholders the following two questions at the first evaluation meeting with them:

• What would you like to be able to do or see as a consequence of this evaluation?
• Why can't you do or see those things right now?

Having received some kind of response (although see #9), do I prepare theories of change for my evaluations and (crucially) do it in a way that allows the theory of change to be challenged and debated? The short answer is: *whenever I can*. When I do, I tend to use models and ideas from the systems field rather than boilerplate theories of change. However, I think that is a weak response. I still feel it places me firmly in the position of being part of the problem rather than some part of the solution.

What I think might resolve this issue, for me and others, is a more advanced discussion about theories of learning and action. Most of those I have seen used

in evaluation are very crude, mostly assuming that in all cases people respond rationally to specific information. Clearly, that is nonsense, yet we still construct theories of change with that mechanistic model of behaviour in the forefront. (Indeed, we still design many evaluations as if that were the case.) My colleagues in the learning field have told me that there have been some substantial advances in understanding cognition in the past few years, which if applied to evaluation would radically change what we do and how we do it. Maybe in a few years we will be incorporating them into our understanding of how to construct theories of change—and maybe we get to close the loop between theories of change and our commitment to evaluation consequences. Of course, we have to acknowledge that at times it is actually unsafe to discuss the consequences of evaluation—we must not be dogmatic about this.

References

Bitar, K. (2022). *Theory of evaluation (ToE): A practitioner's guide to constructing a 'theory of evaluation' and developing evaluative knowledge products*. Retrieved from https://bit.ly/3BOwDrX

Hansen, M., Alkin, M. C., & Wallace, T. L. (2013). Depicting the logic of three evaluation theories. *Evaluation and Program Planning*, *38*, 34–43. https://escholarship.org/uc/item/30c4j0wf

Williams, B. (n.d.). *Qualitative data analysis*. Retrieved from www.bobwilliams.co.nz/ewExternalFiles/analysis2.pdf

Part 6

Developing theories of change that reflect multiple stakeholder perspectives

22 Participatory explorations of alternative futures

Using narrative theories of change

Rick Davies, Lara Mani, and Tom Hobson

Origins

ParEvo is a web application that enables the participatory exploration of alternative futures, including the kinds of desired and expected futures evaluators encounter in theories of change. The design of the process embodied in the ParEvo app has its origins in doctoral fieldwork on organisational learning within an NGO working in Bangladesh during the 1990s (Davies, 1998). That research had its basis in an evolutionary epistemology (Campbell, 1960; Campbell et al., 1987; Cziko, 1997) where evolutionary processes were seen as a form of learning, which could be observed at multiple spatial and temporal scales. One of the practical outcomes of that research was the development of a form of participatory impact monitoring known as Most Significant Change, which involves the reiterated collection and selection of stories of change (Davies & Dart, 2005; Ohkubo et al., 2022; Tonkin et al., 2021).

At the end of that research, it was noted that:

> Other potential applications of the evolutionary algorithm have been identified. One is the participatory development of past project histories (or future project plans) in the form of evolving branching structures of narrative. Participants would choose which branch-end to add a next step in the story. New branches would emerge where an existing branch was added to by more than one participant. Branches would die out where no new additions were placed there by participants. Such constructions could capture something of the contentious nature of history and the way in which it is socially constructed. (Davies, 1998, p. 337)

This possibility was subsequently tested out in a classroom setting in 1996, but it then lay dormant until 2019, when an opportunity arose to develop a web application that enabled people to participate in the process, regardless of location.

Relevance

ParEvo was designed primarily as a means of collaboratively exploring alternative futures and alternative histories. It was not designed specifically for

use by evaluators, nor was it intended to replace existing approaches to the development and use of a theory of change. However, it has some commonalities and differences that could enrich the design and use of theories of change as seen by evaluators. As with many forms of evaluation practice, it involves a participatory process, one which is usable in both face-to-face and online settings. Different stakeholders can be engaged in distinct roles, both in the construction and evaluation of narratives about alternative futures. Its operation online is also not unique, with many concurrent experiments underway with different forms of web-enabled collaboration in the construction and evaluation of theories of change (Montague et al., Essay 25; Macfarlan, 2020).

There are also some major differences. Firstly, the text of storylines developed during a ParEvo exercise provides what Clifford Geertz (1973) described as 'thick description', in contrast to what is in effect skeletal and decontextualised descriptions often provided by diagrammatic representations of theories of change (Davies, 2018). Secondly, ParEvo exercises generate a diversity of possible futures, which vary in their desirability and likelihood, in comparison to the more singular, expected, and optimistic outcomes of many theories of change. This diversity of futures provides the potential for more adaptive programming of interventions than a theory of change; its focus on the desired and expected can make it more akin to a form of 'blueprint planning' (Rondinelli, 1993). Given these characteristics, a ParEvo exercise is likely to be of most use in situations where programme designers are faced with environments that are notably complex and unpredictable.

Use to date

As of late 2021, 13 different ParEvo exercises have been completed. Participants have included school students; volunteers recruited from a monitoring and evaluation community of practice-paid adult, university-educated crowdsourced participants; staff from a UK development aid think tank; UN volunteers; UN agency staff members; and internationally recognised experts in particular fields. Futures explored include post-Brexit Britain, climate change post-COP26,[1] a five-year corporate strategy, post-Trump USA, and governance of biotechnology research. Histories have explored agricultural development project implementation, UN volunteer experiences, and gender policy implementation within a UN agency. Participants in these exercises were both nationally and internationally sourced, and always of mixed gender.

---

1 The 2021 annual meeting of the UN Climate Change Conference.

An example ParEvo exercise

In early 2021, the Centre for the Study of Existential Risk (CSER) at the University of Cambridge, UK, adopted the use of ParEvo to explore future pathways for governing biotechnology research. Following a successful small-scale, in-house trial, a full-scale exercise was undertaken with specialists in the field of biotechnology in November 2021. The exercise had two principal aims: First, to explore possible risks associated with innovations in biotechnology with a group of elite stakeholders in the field and, second, to evaluate the effectiveness of ParEvo as a tool for futures exploration, which may be of wider relevance to futures explorations within the field of global catastrophic and existential risk studies.

Eleven internationally recognised experts from the fields of biotechnology and biosecurity were recruited to be the exercise participants. This number was based on previous successful ParEvo experiences on the same scale, which suggest an optimal number of participants to be between 10 and 12.[1]

The starting point of a ParEvo exercise is a 'seed paragraph' drafted by the facilitator, and akin to the first page of a novel. It sets the scene, usually in the present or nearby, and is best understood as the starting point from which the facilitator wants the emerging storylines to grow. In the CSER exercise this paragraph began: 'In July of 2021, the leaders of the G7 states gathered in Germany for their annual summit…' and ended with 'What follows below are stories of what happened over the next four years'.

The development of a ParEvo exercise proceeds through a series of iterations, each of which represents a defined period of time. In the CSER exercise, each iteration represented a period of six months, and eight iterations were completed—meaning the storylines would eventually cover developments over a period of four years into the future.

The facilitator also provided participants with guidance notes, designed to explain both the aim of the exercise and how participants were expected to engage with the process. The guidance notes are present at the top of each exercise and updated for each new iteration. They are also used to outline the boundaries of what types of content would and would not be acceptable (for example, not science fiction or fantasy). ParEvo also allows for commentary to be made on individual contributions by facilitators or external commentators, to help align participants with the exercise aims and to challenge participants to keep their storylines plausible.

The first iteration involved participants reading the seed paragraph and then drafting a new paragraph, of up to 150 words, to extend the story beginning in a direction of interest to them. This was done by each participant within the same 48-hour time period, allowing participants to contribute across various time zones. Only at the end of the contribution window were all participants able to view each other's contributions, although contributions remained anonymous.

The next iteration then began, as participants read the 11 emerging storylines and were asked to choose one of these that they wanted to extend further by adding their second contribution. It was made clear that they could extend any storylines, either one ending with their own previous contribution or that of any other participant. As expected, contributions described unfolding events that were seen as possible but not necessarily likely or desirable.

The same process, of participants reading contributions made during the previous iterations and then adding their own new contribution to a selected storyline, continued for eight iterations. However, if a given storyline was not extended in the previous iteration, it could not be extended thereafter; that storyline was in effect 'extinct'. The entire process was evolutionary, in that it involved the reiteration of *variation* (viewing the existing range of 11 different storylines), *selection* (of one of those, by each participant), and *reproduction* (the adding of one or more new contributions to selected storylines), thus recreating further diversity.

After the completion of the eighth iteration, participants were asked to evaluate the surviving storylines, by two means. One was using a widget built into the app, where participants selected a storyline that they considered was most likely, least likely, most desirable, and least desirable. The other was via an online survey instrument that asked a wider range of open and closed questions about perceived significant differences between storylines, the most surprising inclusions and omissions in the storylines, perceived optimism of their own and others' storylines, and the perceived ability to influence and be influenced by the described events. In addition to this survey data, multiple datasets were generated by the ParEvo app itself, including the full texts of all storylines (extinct and surviving), data on participants' selection choices, plus the results of keyword searches of all storylines. In the CSER exercise, the 80 contributions by participants created 11 surviving storylines and 20 extinct storylines, ranging in length from 150 to 1,200 words each. Further information was collected through two virtual focus groups with all participants and one-to-one virtual semistructured interviews with individual participants.

Figure 22.1 shows the left side of the user interface, displaying the structure of the storylines that were developed, along with the evaluation judgements made on the one highlighted storyline. The right side of the interface, not shown below, displays the full text of that same highlighted storyline. Users can click on different storylines to see the associated full text.

Subsequently, data from two facets of the exercise were analysed: the responses of the participants and the contents of the storylines. With respect to participation, engagement by participants remained high throughout the period of the exercise (18 days), with the exercise achieving 80 of the possible 88 contributions. During the focus group discussions, numerous participants noted they enjoyed engaging in the exercise, particularly in comparison to other futures exploration exercises (for example, Delphi and horizon scanning).
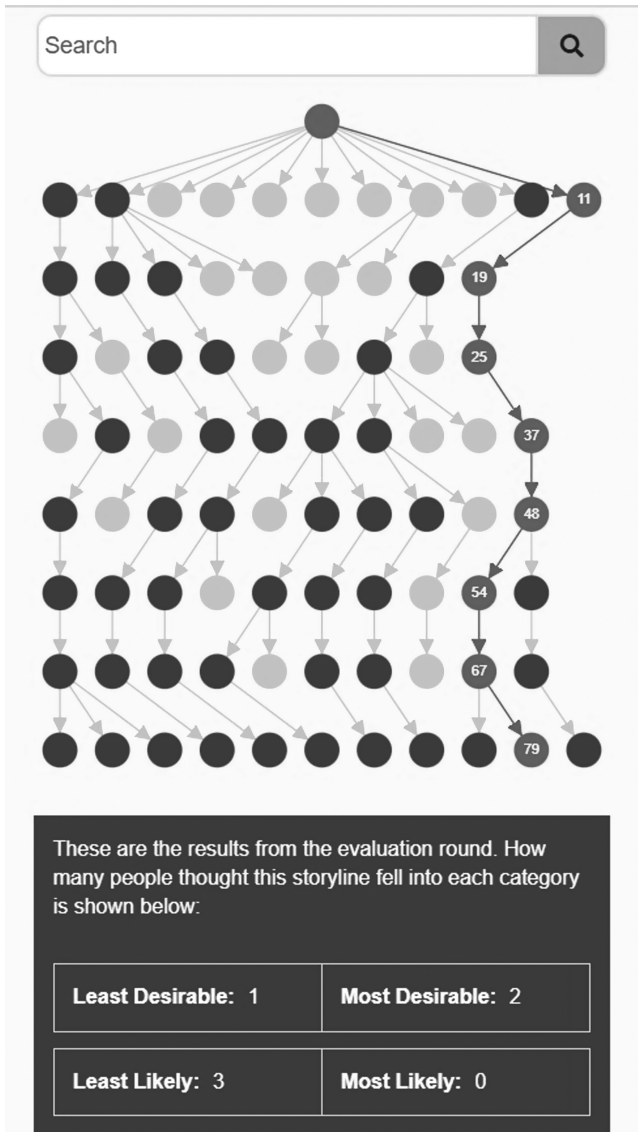
Figure 22.1 The tree structure of the completed eight-iteration CSER exercise showing 20 extinct storylines (grey), 11 surviving storylines (black), and one selected storyline (numbered grey). Each node represents a consecutive paragraph of text.

Participants' contribution behaviour varied on two axes, derived from social network measures known as *outdegree* and *indegree*: the extent to which they added to others' contributions and the extent to which others added to their contributions. Compared to previous ParEvo exercises run by other organisations, more participants were 'isolating' rather than 'bridging', 'following', or 'leading'. That is, participants more often tended to build on their own contributions rather than on others'.

While most participants rated the storylines as more pessimistic than optimistic, most also rated them as realistic. Storylines were diverse when viewed in a scatterplot showing desirability versus likelihood ratings, but there were relatively few in the desirable and likely quadrants. Disagreements were most notable about the desirability of some unlikely storylines. Some of the other differences participants noted between the storylines were:

• The extent to which exogenous shocks were involved
• The role of top-down versus bottom-up processes of negotiation
• The extent to which substantive progress is made on international negotiations versus none at all
• The influence of disinformation/conspiracy theories on negotiation processes.

Common themes seen in the storylines included geopolitical rivalry, major changes in international governance mechanisms, and national-level innovations in research management. Aspects of the content analysis are still ongoing.

Challenges

The CSER exercise, like others before it, generated a wealth of data about participants' views of possible futures, within a realm of concern to them—in this case, the possibilities of effective global governance mechanisms for the management of biotechnology research risks. Analysis of participation behaviour was relatively easy, using simple social network analysis metrics and visualisations. Content analysis was more challenging. Top-down methods finding differences between whole storylines seem more immediately useful compared to bottom-up methods based on coding specific parts of individual contributions, providing a quicker and more useful overview. One such output was a nested classification that progressively differentiated the kinds of storylines based on their outcomes, as evident in the eighth and final iteration. Experimental analyses of the text data also suggest that opportunities exist to construct causal map diagrams of the processes leading to the outcomes, using causal mapping methods developed by Powell and colleagues[2] (Powell & Remnant, 2019). Actor networks, identified through keyword searches for co-occurrences within contributions, are one potential basis for building such maps. However, there are also simpler options,

whereby a ParEvo exercise might simply precede and inform the thinking of the same participants in a theory of change construction workshop.

The design of each ParEvo exercise involves choices, such as numbers and kinds of participants, the participant's roles, the numbers and duration of iterations, and the guidance given. These are likely to affect participants' behaviours and the resulting diversity of storylines. But how best to maximise that diversity is yet to be determined. One finding is clear so far: exercise facilitators need to exercise their 'boundary management' role with care and sensitivity, minimising directive behaviour but not at the expense of the exercise losing direction and coherence. Other means, such as anonymised participation, have already been built in, on the assumption that this will moderate the effects of power imbalances among participants.

Behind the design of ParEvo is a multifaceted view of diversity, as both a cause and a consequence, and as a problematic. Variation is central to evolutionary algorithms and is enabled by the involvement of diverse and independent participants (Dennett, 2019). In complex natural and constructed environments (for example, a ParEvo exercise), diversity is also an outcome. The diversity of storylines generated in a ParEvo exercise is also problematic in as much as they represent multiple possible futures that participants might be able to prevent, enable, mitigate, or capitalise on (if they can be identified). Reflection on these possible responses to the different storylines is an important part of a ParEvo exercise process, and one that needs more attention in future exercises. It is this potential to generate a diversity of practically relevant alternative futures that gives us confidence that the use of ParEvo represents a substantial advance in thinking about theories of change in complex environments.

Diversity is also what distinguishes the results of ParEvo exercises from many diagrammatic theories of change and the outputs of many futures exploration methods. In diagrammatic versions of theories of change reviewed by Davies (2018), some diversity was present in the form of alternate causal pathways but only to a small number of desired outcomes. In futures exploration methods, a common output is a set of four scenarios located in a 2 x 2 matrix structure (Oteros-Rozas et al., 2015). This compares to the 11 surviving and 20 extinct storylines in the CSER exercise. In the surviving storylines alone, there were more than 70 different named (and real) institutions and social structures mentioned, along with references to 29 countries. The analysis of this diversity is manageable, using a range of participatory and other methods.

But is this scalable? Direct participation has been limited to between 10 and 15 people in exercises so far. Limited screen space and participants' cognitive capacity may make it difficult to increase these numbers. Other options are being explored, notably through enabling any number of 'observers' to track an exercise and feed their responses back to the facilitator. In one exercise more than 700 peers of the participants were able to feed their views back through an online survey.

ParEvo in the near future

The latest ParEvo exercise, under design in early 2022, will involve recipients of research grants, funded by an American foundation. Participants will explore alternative possible futures relating to uptake of their research, with a view to then responding to those possibilities, which are expected to vary in terms of their likelihood and desirability, as well as other criteria such as equity. ParEvo exercises will then be repeated each six months thereafter, to enable grantees to update and adapt how they see those possible futures and appropriate responses. In CSER, the most immediate response to the end of the second exercise described in this essay has been the decision to use it again, this time as the backbone of a two-week exchange between participants who met in the 2022 Cambridge Conference on Catastrophic Risk.

Notes

1 However, means do exist for engaging a larger number of participants. These include providing other roles in addition to exercise facilitator and participant, including commentators—on participants' contributions during the exercise, and observers—providing feedback to the facilitator.
2 www.causalmap.app/aboutus

References

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, *67*(6), 380–400. https://doi.org/10.1037/h0040373

Campbell, D. T., Heyes, C. M., & Callebaut, W. G. (1987). Evolutionary epistemology bibliography. In W. Callebaut & R. Pinxten (Eds.), *Evolutionary epistemology: A multiparadigm program* (pp. 405–431). Springer Netherlands. https://doi.org/10.1007/978-94-009-3967-7_20

Cziko, G. (1997). *Without miracles: Universal selection theory and the second Darwinian revolution*. MIT Press.

Davies, R. (1998). Order and diversity: Representing and assisting organisational learning in non-government aid organisations [Doctoral dissertation, University of Wales, Swansea]. http://mande.co.uk/blog/wp-content/uploads/2013/05/thesis.htm

Davies, R. (2018). Representing theories of change: Technical challenges with evaluation consequences. *Journal of Development Effectiveness*, *10*(4), 1–24. https://doi.org/10.1080/19439342.2018.1526202

Davies, R., & Dart, J. (2005). *The 'Most Significant Change' (MSC) technique: A guide to its use*. Retrieved from www.researchgate.net/publication/275409002_The_'Most_Significant_Change'_MSC_Technique_A_Guide_to_Its_Use

Dennett, D. (2019). *Is evolution an algorithmic process?* UWTV video. www.youtube.com/watch?v=ECitgI0x55M

Geertz, C. (1973). *The interpretation of cultures: Selected essays* (1st ed.). Basic Books.

Macfarlan, A. (2020). *Theory of change software*. Better Evaluation, the Global Evaluation Initiative.

Ohkubo, S., Mwaikambo, L., Salem, R. M., Ajijola, L., Nyachae, P., & Sharma, M. K. (2022). Lessons learned from the use of the Most Significant Change technique for adaptive management of complex health interventions. *Global Health: Science and Practice*, *10*(1). https://doi.org/10.9745/GHSP-D-21-00624

Oteros-Rozas, E., Martín-López, B., Daw, T., Bohensky, E., Butler, J., Hill, R., Martin-Ortega, J., Quinlan, A., Ravera, F., Ruiz-Mallén, I., Thyresson, M., Mistry, J., Palomo, I., Peterson, G., Plieninger, T., Waylen, K., Beach, D., Bohnet, I., Hamann, M., & Vilardy, S. (2015). Participatory scenario planning in place-based social-ecological research: Insights and experiences from 23 case studies. *Ecology and Society*, 20(4). https://doi.org/10.5751/ES-07985-200432

Powell, S., & Remnant, F. (2019). *Causal maps: Making connections*. Retrieved from www.causalmap.app

Rondinelli, D. A. (1993). *Development projects as policy experiments: An adaptive approach to development administration*. Routledge.

Tonkin, K., Silver, H., Pimentel, J., Chomat, A. M., Sarmiento, I., Belaid, L., Cockcroft, A., & Andersson, N. (2021). How beneficiaries see complex health interventions: A practice review of the Most Significant Change in ten countries. *Archives of Public Health*, *79*. https://doi.org/10.1186/s13690-021-00536-0

23 Constructing a living theory of change in fluid and volatile environments

Grounded on context, problem, and evidence

Hur Hassnain

Introduction

Theory of change, a key conceptual 'child' of the evaluation profession, is an explicit articulation of how a policy or programme intervention creates an intended result to address a specific problem (adapted from Wholey, 1987, and Rogers, 2014). In simple words, the theory of change outlines and explains how the activities of a development or humanitarian intervention lead to a hierarchy of short-, medium-, and long-term results, with a due consideration of the underlying causal, operational, environmental, and contextual assumptions behind each step of the theory. Hence, it is a theoretical or graphical representation of a collection of hypotheses and assumptions about how the intervention being evaluated works.

A theory of change can be used to better understand the context of operation and how change happens in the given scenario based on assumptions that come through evidence acquired during previous interventions in the same or similar contexts. If it is designed well at the intervention's inception stage, preferably with the help of an evaluation expert, a theory of change can help identify key indicators of change, monitoring, and evaluation questions, as well as data and evidence gaps in a given context.

The 'theory' within a theory of change is specifically difficult to construct if the 'context' is fluid (that is, rapidly changing), complex, or volatile (that is, conflict affected or conflict prone due to different triggers of violence). In these contexts, the relationship between the causes and effects is not linear and plans and strategies often do not work as anticipated. What works today may not work tomorrow. Whilst some activities lead to outcomes, others do not or outcomes (positive or negative) could also emerge in the most unexpected places (Pearl & Mackenzie, 2018). A strong culture of learning, including the documentation of lessons learnt in these contexts, is often lacking. In such situations, causality tends to be messy, multilevel, and multidirectional—as well as unpredictable.

Many of the challenges stem from a lack of understanding of the context and, in turn, a lack of realism about the capacity of external actors to support lasting positive change. Understanding how change happens in a context is essential at the earliest stages of an intervention in these settings. Carrying out robust gender

and conflict analyses and using this knowledge to inform intervention design (theory of change) will help ensure that interventions are feasible, do not exacerbate conflict dynamics and inequality, and, where possible, contribute positively to peace and stability (UK Government Stabilisation Unit, 2019).

This essay discusses why and how evaluators should be engaged at the planning and design stage of an intervention, especially whilst constructing the theory of change and its underlying assumptions based on previously acquired evidence, either from existing literature or from detailed consultations with the stakeholders, including the community members and the various institutions involved. During an evaluation, a theory of change can be reconstructed by the same or a different evaluation team to better illustrate the causal links, hypotheses, and assumptions together with all stakeholders; it should be grounded in the context, problem, and available evidence (if any). Besides enabling an assessment of whether an intervention had a positive, negative, or no effect, this approach may also help to ensure the evaluation process is 'doing no harm' to the already affected population (for example, people living in fragile and conflict-affected contexts).

Theory of change and its assumptions

Assumptions are the key conditions for the realisation of different steps in a theory of change, namely the activities, outputs, and outcomes. They play a big role in theories of change for development of humanitarian policy or programme interventions.

Whilst assumptions are the backbone of any theory of change, in complex situations, they play an even more vital role in understanding and responding to the fluid contexts. The lack of a well-grounded, strong contextual and conflict analysis can lead to interventions causing serious harm to the already affected population; this could be avoided by investing resources in analysing the risks and assumptions better while establishing the causal links for a theory of change (for example, if a conflict is occurring between two rival ethnic groups in a conflict-affected context). If your intervention is designed with an assumption that supporting the most affected group would mean that your organisation is saving lives, you may be missing the fact that by doing this you are directly intervening in the conflict, and this may create negative feelings amongst the rival ethnic group. Increased tensions between rival groups could result in harm to staff in your organisation or again to the affected group you were aiming to help. A similar case happened in Central Asia in the past and some injuries and casualties were recorded.

In fragile and volatile situations, a theory of change could benefit from a series of consultations with different stakeholders to ensure that all assumptions and risks are based on detailed conflict analysis. This activity is important to demonstrate and highlight all the different factors that might lead to change with or without further tensions—even if they are not related to your intervention.

The assumptions in a theory of change can be grouped into four categories: (1) causal assumptions, (2) operational assumptions, (3) environmental assumptions, and (4) contextual assumptions. Although all these categories are important, contextual and operational assumptions play a very important role in contexts affected by fragility and volatility. In the scenario exemplified above, where conflict is present in a given context, an assessment of both the contextual and operational assumptions could lead to better programming and decision making.

Theory of change at the intervention design

Evaluators are rarely included in the initial designing of theories of change for policy or programme interventions. Evaluators work on the already existing theory of change at the time of evaluation, often developed by people with limited or no knowledge of evaluation. In such situations, there is a strong tendency to miss the underlying story of causality or how change may or may not happen over time. This can affect the emphasis on the risks and assumptions in a theory of change.

An evaluator may help the policy or programme staff to understand the causal links between the hierarchy of results together with the causal, operational, environmental, and contextual assumptions based on evidence (if available). Where there is no evidence available, fully or partially, to construct the causal chain, or if there are competing theories especially in a conflict-driven context, an evaluator may work with different groups to either agree on one causal link or work on alternatives in mutual consensus with the people involved. This approach may help an intervention in a fluid and volatile context to construct a learning and sharing environment that is based on a culture of generating solid evidence for decision making by defining better evaluation questions. From there, the final evaluator can be supported to select an appropriate design for the final evaluation and tailor methods and tools accordingly without losing the story of what happened during the lifecycle of the intervention.

Another benefit of engaging an evaluator to develop a theory of change at the design of an intervention would be to identify and capture innovations that led to change with strong evidence generated throughout the life of the intervention. In the words of Michael Quinn Patton, interventions in fragility, conflict, and violence contexts often focus on principles that guide innovation that is needed in 'improvising rapid responses in crisis conditions' in 'complex dynamic environments' (Patton, 2018). This is possible if the theory of change is grounded in the context, the problem for which the intervention was designed to address, and the available evidence.

A theory of change, in fragile situations, could better inform the development of a strong monitoring and evaluation framework, with a mapping of where there is existing data (available from the intervention and/or from previous research

and evaluation) and where strong emphasis could be laid to gather new data and identify weak or no causal links, based on contextual factors and to investigate patterns, for example, if the intervention worked well in any particular sites.

An interesting way of developing a theory of change at the programme design stage could be to identify possible causal links based on different positive or negative scenarios or, in other words, *assumptions* with a group of relevant stakeholders. This is not always possible, but it is a very useful exercise in my experience so far.

Theory of change at the evaluation stage

Rigorous evaluations, including those conducted in fragile contexts, reconstruct the intervention's theory of change at the outset of the evaluation process. Ex-post reconstruction of the theory of change allows the evaluation to assess not only *if* but also *how* and *why* the intervention has or has not had any results. For an evaluation in a moving and volatile context, the theory of change should be reconstructed, if possible, together with all relevant stakeholders, to identify and then test the different types of underlying assumptions. This effort will help in understanding the drivers of conflict during the life of the intervention, how and if the intervention had different impacts on specific identity groups, and how to connect the outcomes and indicators with gender and conflict analyses underpinning the evaluation design (Hassnain et al., 2021).

Understanding the design and implementation of the intervention under evaluation is a key factor in developing an appropriate evaluation design. This understanding is particularly crucial for evaluations conducted in contexts affected by fragility, conflict, and violence, as it enables the testing of assumptions underlying the theory of change about how development interventions affect change, which is in turn important for understanding the results (Gaarder & Annan, 2013). This understanding stems from a deep dive into the contextual analysis either conducted previously by experts or as part of the evaluation exercise conducted by the evaluation team. The evaluation commissioners should ideally refer to the availability of a conflict analysis in the evaluation terms of reference, and this should be highlighted by the evaluation team in the evaluation inception reports, together with any budgetary implications it may have. It is possible that the evaluation team may have to find additional support to conduct a gender and conflict analysis if one is not yet available.

Without a full understanding of the context in which an intervention is operating, it is hard to assess the results and choose appropriate evaluation designs, tools, and methods. This point is demonstrated by the Independent Commission for Aid Impact in its 2018 review of the UK Conflict, Security and Stability Fund (CSSF). It says, 'A theory of change must cover a coherent effort. In the case of the CSSF, this would often include one or more CSSF programmes (ODA and non-ODA), as well as diplomatic or defence engagements and possibly other

UK and wider international efforts. We saw a number of theories of change that consisted only of a few bullet points… Such a superficial ToC is not useful for planning, monitoring, or evaluation purposes' (Independent Commission for Aid Impact, 2018, pp. 18–19).

If it is at all possible, an outline of the theory of change and its underlying assumptions could be discussed with members of an evaluation reference group and different stakeholders during the inception meetings so that the causal pathways are identified together and tested throughout the evaluation in a conflict-sensitive manner. The key points to focus on would be to identify what evidence the intervention was based on and what additional evidence has become available since then to prove the causal links in the given context. In my experience, designing a theory of change with a colour-coding of evidence at different levels of results gives a better picture of whether there were gaps in the evidence that has been used to develop the theory of change or indications that it has changed since being developed. This occurs before the evaluation team revises the theory of change with the help of data gathered during the evaluation exercise.

It is to be noted that the theory of change at the inception stage should be based on existing theories of change and related programme and policy documents (such as intervention logic, logframe, and others), so it speaks to the intervention and can be compared to the change story at the design stage and at the end of the intervention. I have been informed of some cases where evaluators created a new theory of change, totally from scratch. Maybe it is a broader question to ask ourselves: How far can an evaluator go at the end of an intervention in defining 'how change happens?' compared to 'how change happened or did not happen?' Maybe evaluators could construct two theories of change at the end of an evaluation: one discussing 'how change happened or did not happen', and the other discussing how they believe change happens in each context.

A theory of change grounded in context, problem, and evidence and co-designed with relevant stakeholders could help in not only programme and policy design, but also its planning, management, monitoring, and evaluation by facilitating a shared understanding of how a programme is intended to work. In this book, Sebastian Lemire (Essay 24) discusses that programme theories can be a useful tool towards promoting equity and transformative change in an evaluation.

Engage stakeholders with caution!

The COVID-19 pandemic has pushed evaluation practitioners to adapt their ways of working in a variety of ways, just like any other profession. During this time, I have seen and experienced a variety of cases where theories of change were co-designed in hybrid ways, both online and in small, in-person gatherings. I have seen in these meetings that the power resides mostly in the hands of local evaluators and practitioners, promoting capacities in evaluations and further developing confidence, negotiations, and communications skills. In their

essay, Rick Davies, Lara Mani, and Tom Hobson discuss some good ways of co-developing narratives and future scenarios using a web application known as ParEvo[1] instead of a one-consensual view approach.

Although stakeholder engagement is useful when building theories of change, it also provides a safer space to test assumptions and re-engineer the programme in real time. It is to be noted that working effectively in the contexts of fragility and conflict differs fundamentally from practices in conventional settings (Cechvala & Miller, 2020). This applies to connecting with stakeholders—both in person and in online meetings.

Theory of change co-design leads to stakeholders' skills acquisition: negotiation skills, improved interactions with authority figures around negotiating boundaries, anger management, risk acknowledgment, and increasing risk aversion. Theory of change co-design also highlights programme discrepancy, prompting re-deployment of human resources.

See Table 23.1 for a quick scan of how constructing theories of change with stakeholders differs for fluid and volatile environments, such as contexts affected by fragility, conflict, and violence.

In such settings, a different approach to stakeholder engagement and community relations is required, especially to tackle questions of potential bias, trust, the sensitive nature of discussions, the risk of violence, and so on. I remember a case in Swat Valley, Pakistan, in a Taliban-affected community, where the focus group discussions conducted as part of an evaluation escalated the conflict. These types of situations must be avoided with detailed conflict analysis and by working with representatives who know the context and are trusted by the communities, including marginalised ethnic groups and vulnerable populations. Local enumerators for data collection should be hired carefully and in ways that diverse community members perceive to be fair, equitable, and transparent.

It is extremely important to maintain constructive relationships throughout and proper buy-in of the evaluation by the community and the government. In an evaluation conducted in Pakistan in 2017, we had to contact the relevant government departments before carrying out data collection in rural areas. The evaluation team did not get permission to use mobile and tablet devices with GPS services for data collection. In the North Kivu region of the Democratic Republic of Congo in 2015, the community did not agree to a child-led data collection activity from their peers and the evaluation team had to hire adults for this activity.

It is strictly not recommended to engage stakeholders in designing your theory of change in contexts of conflict and fragility without establishing a sold conflict-sensitive approach. The questions around whether to continue operating, what level of engagement to have—if any at all—and with what intent to continue working are all inescapable in such settings (Scudder, 2017).

---

1 For more information, see parevo.org.

Table 23.1 Constructing theories of change in different contexts

Themes	Normal contexts	Fluid and volatile contexts
Theory of change and assumptions	The causal links, as well as operational and environmental assumptions, are often well established to create a theory of change or there is at least some evidence of 'what works'.	Mostly there is no evidence of what works or how 'change' happens in a given environment. It makes it hard to establish causal assumptions. Contexts are unpredictable, so it is hard to establish operational and contextual assumptions.
Theory of change and intervention design	There is often enough time to identify and include multiple stakeholders and come to an agreement on different aspects of theory of change formulation.	The interventions and their designs are often rushed due to emergency situations, so it is hard to identify and engage stakeholders, even thematic experts and evaluators. This may result in unavailability or a poorly designed, nonconflict-sensitive theory of change for the intervention. Conflicting opinions and dealing with bias are sometimes additional barriers to getting people to agree to a common definition of *change*, more common in 'peacebuilding' interventions where *peace* means different things to different people.
Theory of change and evaluation	Quality evaluators are available as per the context and thematic areas of evaluations. Reconstruction of a theory of change can happen based on contextual knowledge, good documentation, and engagement with stakeholders. Stakeholder maps are available, and they are accessible to talk.	It is difficult to identify and recruit quality evaluators and enumerators. Stakeholders are often not identified or trusted due to different biases at various levels. The population, authorities, or institutions may be moving or rapidly changing (for example, the situation in Afghanistan over the last few years). It is difficult to establish a clear causal connection between 'what has changed/is changing' and because of whom.

Conclusion

A well-articulated theory of change that is based on robust assumptions of 'what works' helps to strengthen the intervention design and define realistic results for better monitoring and evaluation. In a situation that is rather fluid,

and volatile, or where primary data gathering may pose serious risks, creating and using a theory of change may provide an opportunity to ensure integration and coordination with other stakeholders or interventions and to build a stronger evidence base. This is fundamentally different from a conventional situation or approach.

Theory of change is the foundation of any policy or programme, and hence it is suggested to engage evaluators from the very outset to design a robust and evidence-informed theory of change together with key stakeholders and thematic experts. This may require a mapping of existing evidence of what works, together with possible alternative scenarios of what could go wrong, to avoid causing harm to the already affected population and discussions about assumptions and risks with different groups of people. This approach would help in better defining the monitoring and evaluation framework for more focused and targeted conflict-sensitive data gathering. In addition, it would ease the defining of robust evaluation questions and a better evaluation framework for the mid-term, final, and post-intervention evaluations.

Evaluation exercises, especially in fluidity and volatility, would benefit from reconstructing the intervention's theory of change at the outset of the evaluation process. If a theory of change does not exist, an evaluation team could spend some time and energy in designing one, with key stakeholders, based on relevant programme and policy documents and a strong contextual, gender, and conflict analysis. A gender- and conflict-sensitive approach in evaluation would contribute to an improved understanding of the drivers of conflict and strengthen efforts to 'do no harm' to the already affected population.

References

Cechvala, S., & Miller, B. (2020). *Memorandum on stakeholder engagement in conflict contexts*. CDA Collaborative. Retrieved from www.cdacollaborative.org/publication/memorandum-on-stakeholder-engagement-in-conflict-contexts/

Gaarder, M., & Annan, J. (2013). *Impact evaluation of conflict prevention and peacebuilding interventions.* Policy Research Working Paper No. 6496. World Bank. Retrieved from https://openknowledge.worldbank.org/handle/10986/15856

Hassnain, H., Kelly, L., & Somma, S. (Eds.). (2021). *Evaluation in contexts of fragility, conflict and violence: Guidance from global evaluation practitioners*. International Development Evaluation Association. Retrieved from https://ideas-global.org/wp-content/uploads/2021/06/EvalFCV-Guide-web-A4-HR.pdf

Independent Commission for Aid Impact. (2018). *The conflict, stability and security fund's aid spending: A performance review*. Retrieved from https://icai.independent.gov.uk/wp-content/uploads/The-CSSFs-aid-spending-ICAI-review.pdf

Patton, M. Q. (2018). Evaluation science. *American Journal of Evaluation*, *39*(2), 183–200.

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect* (1st ed.). Basic Books.

Rogers, P. (2014). Theory of change. *Methodological briefs: Impact evaluation no. 2*. UNICEF Office of Research.

Scudder, D. (2017). Stakeholder engagement in conflict-affected areas: A perspective on recent practice [Doctoral dissertation, University of Tokyo]. Retrieved from www .soka.ac.jp/files/ja/20170419_142533.pdf

UK Government Stabilisation Unit. (2019). *Guide to MEL in stabilisation and conflict*. Retrieved from www.gov.uk/government/publications/monitoring-evaluation-and -learning-mel-in-conflict-and-stabilisation-settings-a-guidance-note

Wholey, J. (1987). Evaluability assessment: Developing program theory. In L. Bickman (Ed.), *Using program theory in evaluation* (pp. 77–92). New Directions for Program Evaluation, No. 33. Jossey-Bass.

24 Flipping the script on programme theories

Advancing towards transformative theories of change

Sebastian Lemire

Introduction

Programme theories are widely used in evaluation. In basic terms, a programme theory is a description of the way in which a programme is intended to bring about a desired set of outcomes—'a plausible and sensible model of how a programme is supposed to work'. Program theories accomplish this by making explicit the underlying assumptions about how specific components of the policy, programme, or project are causally connected to one another and the intended change. In this way, a programme theory basically represents a shared set of assumptions about how things work.

There are many benefits to using programme theories. If designed and implemented well, programme theories support better programme design, management, and evaluation by facilitating a shared understanding of how a programme is intended to work. Evaluators use programme theories for communicating what the programme is and what it is trying to accomplish to external stakeholders, to assess implementation fidelity, and to guide and structure data collection for impact evaluations, among other purposes. To be sure, the role and purpose of programme theories in evaluation continue to evolve.

The purpose of this essay is to make the case that programme theories can be a useful tool toward promoting equity and transformative change in an evaluation (Lemire et al., 2019). Toward this end, I first provide a brief discussion of the issue of biases in programme theories and how this issue necessitates rethinking the current foundation of programme theories. Rooted in Leiderman's (2005) call for 'flipping the script' on evaluation practice, the notion of transformative theories of change is introduced. Toward advancing transformative theories of change, I then discuss the foundational aspects of our current practice that will have to be transformed. This includes rethinking the purpose of developing theories of change (the *why*), *how* and *for whom* we develop theories of change, and what we focus on in our theories of change (the *what*). In conclusion, I briefly discuss the implications of transformative theories of change for reflective practice.

Before advancing the present essay any further, a brief note on terminology is required. As noted, and correctly so, by an astute editor of the present book,

the terms *equity* and *transformative change* are often bandied about. For the purpose of this essay, I use *equity* to reflect fairness and inclusiveness in people's ability to thrive and achieve the best possible life outcomes. This use of the term recognizes that we do not all start from the same place or are afforded the same opportunities; this is why making adjustments to these imbalances is called for. In extension, I define *transformative chang*e as dismantling the underlying root causes and structural barriers maintaining these imbalances. This entails shifting power dynamics; reducing disparity, exclusion, and discrimination; and increasing the autonomy and voice of people who have been marginalized or excluded based on race, ethnicity, gender, ability, sexual orientation, and other dimensions. This use of *transformative change* assumes that the change promotes equity by reducing barriers and ensuring fairness and inclusion in the distribution of and access to quality education and health care, economic and social wealth, among other drivers of inclusive public good.

Flipping the script on programme theories

The issue of bias is nothing new in evaluation. As Sally Leiderman (2005, p. 91) reminds us:

> Our ways of knowing and our trust in others' ways of knowing depend, today at least, in large part on the way our professional and personal lives have been shaped by racism, other methods of oppression, white privilege and access to power.

As human beings, we tend to project our current mindset and assumptions onto the past and future. These biases influence the questions we choose to ask, the information we trust, which findings we decide are important or unimportant, and how we make meaning of results—effectively reinforcing advantages for some and disadvantages for others (Dean-Coffey, 2018). This is certainly the case with programme theories that are heavily influenced by social norms, biases, and privilege, all of which are internalized at the individual, organizational, and systemic level (Dean-Coffey, 2018).

This, of course, is highly problematic as programme theories serve to structure, guide, and focus programme development, planning, and evaluation. To make matters worse, and as discussed later in the essay, many common practices around programme theory development reinforce these biases. These are practices that will have to be undone if we are to promote transformative theories of change.

Toward this end, I will in what follows identify what I take to be some fundamental aspects of programme theories that will have to be changed. Specifically, I will consider some of the common practices that are problematic and propose strategies for changing these. These are not exhaustive of all the issues and

strategies that should be considered but more illustrative of the kind of reflective practice that I think is called for if we are to advance transformative theories of change in evaluation. I will structure my reflections around why, for whom and how, and what we focus on in our programme theories.

Why we develop theories of change

We often frame theories of change as a product and process that supports a shared understanding of what the programme is—how it is intended to generate change. The primary purpose—or the *why*—of the theory of change is to support programme planning, implementation, evaluation, and monitoring. This primarily benefits evaluators, programme staff, and to some extent programme funders.

If we are to promote transformative theories of change, the first essential step is that we center the purpose of our theories of change on transformative change. As stated earlier in this essay, this entails shifting power dynamics; reducing disparity, exclusion, and discrimination; and increasing the autonomy and voice of people who have been marginalized or excluded based on race, ethnicity, gender, ability, sexual orientation, and other dimensions. The primary benefit of transformative theories of change is for those who are most likely to be affected by the programme.

How and for whom we develop theories of change

A second and equally important step is to reconsider *how* and *by whom* programme theories are developed. Program theories are most often developed by evaluators based on programme documents and, if developed in a participatory manner, based on input from programme staff and funders (often referred to as stakeholders). The participatory process is facilitated by the evaluator. This practice is problematic because the resultant theory of change is unlikely to reflect the diverse values and experiences of the people and communities most likely to be influenced by the programme.

If we are to promote transformative theories of change, we must in the development of the theory of change shift power to those that are most impacted by the programme—the true stakeholders. This entails being more inclusive and paying more attention to the different subgroups of the programme participants, awarding particular attention to people who have been marginalized or excluded for various conditions.

It is important to note that this is not just about including a broader range of stakeholders in the theory of change development. By simply including people most impacted by the programme, especially people that are traditionally marginalized and excluded, along with funders and programme staff, existing power dynamics are likely to remain and influence the theory of change. Careful consideration should be awarded to how and with what purpose people are

included. Fortunately, there is a broad—and still broadening—range of cultur-ally responsive methodologies and participant-driven processes for addressing power dynamics that are relevant in this regard. As just one example, the Racial Equity Tools (n.d.) website provides multiple strategies and tools for the equita-ble development of theories of change.

In a recent talk on this topic, several attendees raised concerns about the potential time and resources needed for this broader engagement. Is the idea to facilitate separate workshops for funders and stakeholders? What if the funders and stakeholders substantially differ on the theories of change? Is this feasible with limited budgets and time? To be sure, purposeful engagement of a broader range of stakeholders has significant implications for how theories of change are developed. That being said, and while recognizing the common constraints of budgets and time, transformative theories of change will necessarily need to reflect the diverse values and experiences of the people and communities most likely to be influenced by the programme. Can we really afford not to engage the people and communities for which the change is intended? While flipping the script on our current practices can be difficult, I cannot help but wonder if we can truly promote transformative change without making this commitment—a topic I return to in the concluding part of this essay.

What we focus on in our theories of change

The third and final proposed shift relates to what is included in our programme theories. More specifically, I will focus on the need to push our practice beyond centering programme theories on surface-level and positive outcomes, linear outcome trajectories, and the inclusion of traditional context factors.

Advancing beyond surface-level outcomes

Common practice in programme theories is to focus on the immediate, surface-level outcomes of the programme. If we are to promote transformative theo-ries of change, we will need to reach beyond traditional outcomes to focus on outcomes that speak directly to the root causes of inequities. Root causes can be defined as 'the underlying factors that create social issues and make those issues likely to persist even though a programme may be in place to alleviate more surface-level needs of individuals and communities' (Building Movement Project, 2013). As just one example, instead of focusing the programme theory on increased student academic achievement—a common outcome in educa-tion—the transformative theory of change focuses on gaps in student academic achievement or, even better, on the underlying, structural reasons for the gaps in student academic achievement. Leiderman (2010) provides relevant examples of racial equity outcomes, transformative public policy changes, and transforma-tive changes in narratives about race, among other outcomes to consider when developing transformative theories of change.

Advancing beyond intended, positive outcomes

Another practice that needs to be changed is the common focus of programme theories on a select set of intended and positive outcomes (that is, outcomes that represent a beneficial change to programme participants and other stakeholders). While programmes want to do good, we cannot assume that programmes always do good. Many programmes have unintended consequences that may or may not be positive. Examples include collateral side effects (adverse spillover effects of the programme), paradoxical or counterproductive effects (opposite of the intended programme effect), inequitable effects (unfair differences across programme participants), and null effects (ineffective programmes).

Speaking directly to this issue, Bonell et al. (2015) advocate, and correctly so, for increased attention to harmful consequences when developing theories of change. Specifically, the authors propose:

- Reflecting on the unintended consequences that can potentially emerge from the interaction between the programme and the context within which it is implemented
- Comparing the theory of change for the programme with the theories of change, intervention descriptions, and/or process evaluations of similar programmes to identify harmful effects
- Consulting with relevant stakeholders to identify how programme activities and mechanisms might be derailed, leading to harmful consequences.

These are useful strategies. In addition to the suggestions above, increased attention must be paid to the ways in which harmful consequences affect different stakeholder groups and likely fall most heavily on the most disenfranchised. Thinking carefully about who is likely to be affected by the adverse consequences is fundamentally a question of equity. Capturing adverse consequences in this way provides a more complete understanding of how programmes work, serves well to inform future programme designs, and over time supports the development of programmes that are less likely to cause harm.

Advancing beyond linear outcome trajectories

Another aspect of programme theories that needs to be reconsidered relates to outcome trajectories. Program theories tend to frame programme participants as a homogenous group by depicting uniform activities, mechanisms, and outcomes for all participants. This is also reflected in the common assumption that change is linear, emerging through a uniform short-, medium-, and long-term outcome trajectory across all programme participants. These assumptions are likely flawed and misleading. Different subgroups of programme participants will likely experience and respond to the programme in different ways, resulting

in different outcome trajectories. For this reason, careful thought should be awarded to the different types of outcome trajectories that could be relevant to depict different subgroups in the programme theory.

Advancing beyond traditional context factors

Finally, the context factors included in our programme theories need to be reconsidered. Current practice tends to focus on institutional and social aspects of the setting within which the programme is embedded. To advance transformative theories of change, we must focus our context factors, for example, on structural racism—a shorthand term for the many systemic factors that produce and sustain racial inequities (Lawrence et al., 2009). As Lawrence et al. (2009, p. 7) observe, 'structural racism is a very complex, dynamic system with interlinked social, political, and economic components'. These include (but are not limited to) public policies and laws, social and institutional practices, as well as cultural norms and representations. For this reason, transformative theories of change need to capture how interconnections between the programme and existing policies, social and institutional practices, and cultural representations and narratives reinforce inequities. By capturing and closely examining these in our transformative theories of change, we will be better positioned to see how race, privilege, and disadvantage remain interconnected with and influence the programme to be evaluated (Lawrence et al., 2009). The Aspen Institute has developed a very useful step-by-step guide on how to incorporate structural racism analysis in programme theory development (Lawrence et al., 2009).

Concluding thoughts: From good plumbers to reflective practitioners

In his work on revisiting the foundations of evaluation, Schwandt (2015, p. 144) advocates for evaluators to cultivate 'a life of the mind for practice'—integrating theoretical understanding and practical reasoning in developing sound judgments about how to design and conduct our evaluations. Schwandt (2015) likens this to a shift from being good plumbers (technicians) to being reflective practitioners. This entails reaching beyond the technical and practical aspects of our practice to engage with and be guided by our ethical and moral commitments. For Schwandt, cultivating a life of the mind for practice is about cultivating discretionary judgment, critical thinking, and reflection—reaching beyond surface-level solutions to do evaluation work that best serves the public good. We need to be reflective plumbers.

This proposed shift toward reflective practice is, of course, a choice evaluators will have to make. In reflecting on racial justice and evaluation, Leiderman (2010, p. 32) observes:

> It seems obvious that we would turn this lens inward to the theories of change and logic models that we use to evaluate work with racial goals.

> But our experience as evaluators suggests this takes real intention and some courage on everyone's part.

While recognizing that flipping the script on our current practices can be difficult, I cannot help but wonder if we can truly be reflective practitioners without making this commitment. Holding on to the imagery of plumbers, consider the following example. Imagine a long-time client reaching out to you—the evaluator and trusted advisor—about a water leak in her office. As your distressed client goes on to explain, a water pipe has burst, steadily dripping water all over the floor. As a trusted advisor, you now have two options.

One option is to advise your client to get in her car and drive to the nearest Home Depot to purchase one of their signature $4.99 orange plastic buckets. Placing the bucket under the broken pipe will immediately remedy the situation, allowing the floor to dry. Sure, the bucket will over time have to be emptied, and the tension in the pipe causing the leak will likely result in other leaks. More buckets will follow. That being said, option one is an easy-to-implement and swift solution—at least from a purely technical and practical perspective.

Another option is to advise the client to take on and tackle the root cause—the leaking pipe. This involves calling a plumber—a reflective plumber—who will examine the leak, identify the underlying reasons for the leak, and eventually fix the leak and any other issues with the pipes. This, of course, will come at a higher upfront cost and require more of everyone involved. That being said, option two is a more permanent solution that speaks directly to the root cause of the problem.

As evaluators, we face leaks all the time—metaphorically speaking, of course. In the end, and as the trusted advisor of our clients, the choice is up to us. Will we be the plumber that advises the client to purchase a plastic bucket, or will we be the reflective practitioner that challenges our clients to fix the broken pipes? The choice is ours.

References

Bonell, C., Jamal, F., Melendez-Torres, G. J., & Cummins, S. (2015). 'Dark logic': Theorising the harmful effects of public health interventions. *Journal of Epidemiology and Community Health*, *69*(1), 95–98.

Building Movement Project. (2013). *Discussion guide: Theory of social change*. Retrieved from https://buildingmovement.org/tools/theory-of-social-change/

Dean-Coffey, J. (2018). What's race got to do with it? Equity and philanthropic evaluation practice. *American Journal of Evaluation*, *39*(4), 527–542.

Lawrence, K., Anderson, A. A., Susi, G., Sutton, S., Kubisch, A. C., & Codrington, R. (2009). *Constructing a racial equity theory of change: A practical guide for designing strategies to close chronic racial outcome gaps.* The Aspen Institute Roundtable on Community Change. Retrieved from www.aspeninstitute.org/wp-content/uploads/files/content/images/Roundtable%20on%20Community%20Change%20RETOC.pdf

Leiderman, S. (2005). Doing evaluation differently. In M. Potapchuk, S. Leiderman, D. Bivens, & B. Major (Eds.), *Flipping the script: White privilege and community building*. MP Associates and the Center for Assessment and Policy Development.

Leiderman, S. (2010). How do we know it when we see it? *Marking Progress: Movement Toward Racial Justice*, *3*, 31–36. https://akonadi.org/wp-content/uploads/2016/07/Marking-Progress-Movement-Toward-Racial-Justice-PRE-Critical-Issues-Forum-Volume-3-1-1.pdf

Lemire, S., Whynot, J., & Montague, S. (2019). How we model matters: A manifesto for the next generation of program theorizing. *Canadian Journal of Program Evaluation*, *33*(3), 414–433.

Racial Equity Tools. (n.d.). *Theory behind the work*. Retrieved from www.racialequitytools.org/resources/evaluate/defining-the-work/theory-behind-the-work

Schwandt, T. A. (2015). *Evaluation foundations revisited—Cultivating a life of the mind for practice*. Stanford University Press.

25 Theory-based evaluation approaches can enable online project success

Steve Montague, Heloise Emdon, and Eva Grabinski

Introduction

Theory-based evaluation provides a critical approach for clarifying the focus, scope, and methodology of an evaluation project while building shared understanding and trust between the evaluators and clients. The application of theory-based approaches in evaluation can help teams work remotely on evaluation projects through online co-development of artefacts, such as theories of change, that facilitate the communication and capturing of key aspects pertinent to an evaluation. This essay considers how theory-based evaluation can facilitate remote collaboration between evaluators and clients, how theory-based evaluation involves an iterative process for obtaining conceptual clarity about the evaluation focus and context, and how theory-based evaluation can be applied to deliver learnings throughout an evaluation project to drive programme improvements.

Two case studies are described that involve practicum projects forming part of the 16-month online postgraduate diploma in policy and programme evaluation (DPPE) at Carleton University that includes students from across the globe. The practicum project serves as a focus for the theoretical curriculum and involves client engagements, evaluation design, planning, execution of evaluation research, and delivering findings to clients.

All course and practicum work typically takes place online. This was true prior to the COVID-19 pandemic; however, it is still notable that the projects described took place during this time period since in at least one of the cases this created significant extra constraints.

Two cases are described and analyzed for key learning. They shall be referred to as the *historically disadvantaged institution* and the *community health care* cases, respectively.

Background

The historically disadvantaged institution case

Early exchanges on a project to evaluate initiatives to improve academic literacy and academic writing at Mangosuthu University of Technology (MUT) started

shortly before the pandemic travel restrictions. MUT is a historically disadvantaged institution, South Africa's only university solely located in a township. In late March 2020, the country instituted one of the hardest lockdowns in the world, and the university closed for six weeks. Staff were ill-equipped to work from home, and most MUT students had limited access to the internet or laptops. During these crucial early months, when the interviews with clients were cut short, the evaluation students went back to the recorded interviews and the few programme documents they had collected and transcribed and coded them using a context-intervention-mechanism-outcome approach, which assisted the team members to distinguish between the sticky concepts of context and the intervention (Punton et al., 2020). This enabled them to distill the generative mechanisms of change that would enable students to become confident in academic concepts and writing. Using research to prepare for truncated consultations before evaluation planning, the student team was able to convert theories into testable hypotheses to undertake a formative evaluation of the programming of the Academic Language and Literacy Unit (ALLU) and the establishment of a new writing center that was initially conceived of as an extension of the ALLU.

To develop a robust theory of change, the evaluation team used a construct of capabilities, opportunities, and motivation (Mayne, 2015, 2017; Michie et al., 2011); this helps the evaluator describe the behavior change and articulate assumptions the intervention was making to achieve these behavior outcomes. Using the Kirkpatrick Model as a reference, the practicum team was able to draft a visual theory of change. Over several iterations and discussions with the client and with DPPE supervisors, it became clear that the focus on students alone did not sufficiently reflect important considerations of how the engagement with these students was achieved. Student engagement relied on an enabling environment where the faculty endorsed participation in the literacy and writing interventions (Baba et al., 2020). Both the programme theory and the generative change analysis helped the practicum team develop a simplified impact pathway (outcome chain) (Funnell & Rogers, 2011).

Applying the Purposeful Program Theory (Funnell & Rogers, 2011) to distinguish interventions (theories of actions) from the sequence of intended learning outcomes (theories of change) and the several if-then assumptions (programme factors and nonprogramme factors) that were distilled from the transcripts, it became apparent there were nested theories of change—one relying on the faculty lecturers to engage students and the other on the programme interventions for the students. This helped scope the practicum evaluation on the faculty lecturers to assess whether the environment was conducive for literacy programming.

The community health care case

As part of community health care, the Early Years Breastfeeding Support (EYBS) Project was conducted to evaluate access by expectant and new mothers

to breastfeeding support services provided by Pinecrest-Queensway Community Health Centre (PQCHC) in Ottawa, Canada. Collaboration between the evaluators and client team was conducted entirely online due to the COVID-19 pandemic. A critical success factor that supported online collaboration was the application of a theory-based approach, which involved a theory of change and a socioecological model shared among the evaluators and client team.

The theory of change proved to be a critical tool for obtaining a common understanding among the evaluators and client team about the primary focus of the evaluation. Importantly, the theory of change was foundational for developing the evaluation plan, including articulating the scope, focus, assumptions, and research questions and methods. For example, the theory of change stipulated that *mothers become aware that the services exist*. The related assumption is that *the client effectively promotes these services to mothers*. In turn, this supported the identification of key research questions and methods applicable to evaluating promotional activities and outputs in reaching mothers as part of access to the services.

A socioecological model was used based on the prevalence of this approach in the health care sector and based on the description of socioecological theory in evaluation (Funnell & Rogers, 2011). The evaluation team ascertained the use of socioecological models in health care via resources such as the Centers for Disease Control and Prevention (2021). The socioecological model used in the EYBS project drew on design thinking from human-centered design (IDEO, 2015). The application of a socioecological model provided a picture of expectant and new mothers as the users of the services; the communities around mothers that influence breastfeeding decisions (for example, partners, friends, peers, and other family members); the health care services and resources available to mothers for breastfeeding support; and the institutions, or systems and policies, that can influence the breastfeeding decisions of mothers.

Clarity through an iterative theory-building process

The building of theories and models to guide a theory-based approach to evaluation entails an iterative process of discovery for both the evaluators and clients. The value of such a process has been noted by Mulcahy, Hassnain and Lemire in this compendium. It is through an iterative process of refining artefacts, such as a theory of change and a logic model, that clarity emerges about the focus and scope of an evaluation.

The historically disadvantaged institution case

In their quest for 'theory gleaning' (Manzano, 2016), the practicum team started to discern subtle differences between the ontological and epistemological foundations of the ALLU and the writing center. Further interrogation of the

transcripts led to their realization that while the goals of the two initiatives were apparently the same, their causal pathways were different. The ALLU and the writing center targeted different stakeholders and worked through different inter-linked mechanisms to achieve the intended impacts. Given the differing causal pathways underlying these integrated student support services, this required developing a deeper understanding of the contradictory programme approaches that needed to be revealed carefully to the client (Dahler-Larsen, 2018).

When the lockdown was eased and regular meetings with evaluators resumed online, the team suggested undertaking a summary realist synthesis to pick through the literature and find evidence-informed case studies of various approaches to 'academic literacy' and 'writing support' to understand how these interventions work in different settings, for different groups, before determining the data they wanted to collect for 'reality testing' (Pawson et al., 2004) and for designing the interview instruments to understand what decisions were made in the delivery of the intervention. The client was extremely interested in this approach and joined the team in the meta-analysis of evidence-based studies.

After coding some 40 journal articles, the evaluators discovered at least three schools of thought relating to the relatively young field of genre studies (Hyon, 1996), each with their own epistemologies. The ALLU's pedagogical under-pinnings were found to come from the field of new literacies, which emerged from English as a Second Language studies in which students needed to develop different kinds of literacies (academic, information, mathematical, and others). Academic literacy is not discipline specific, and it is delivered in an add-on fashion, both in terms of the context of the student and also in terms of the con-text of the discipline. Academic literacy, which refers to language and linguistic capacities, has been sharply critiqued for creating 'decontextualized learners'—students treated as if they have no social context—and an 'autonomous model of literacy' in which the use of language is seen as a set of neutral skills (Boughey, 2016).

By contrast, the pedagogy in the writing center was informed by socially oriented theories of learning also referred to as the *new rhetoric genre theories*, which emanated from North America with an ethnographic focus (Artemeva, 2008). This theoretical lens is derived from the fields of rhetoric and rhetorical genre theory. It perceived all writers to be novices in terms of learning how to use the genre in which they were learning to write.

The contrasting perspectives were subtle initially, but eventually—as the practicum group gathered more data from key informant stakeholders—the team realized that contradictions in the approaches ran deeply through the pro-gramme theory with diverging ideas of how the programme was supposed to work. The teams' triangulation of methods drew them further into testing the implicit assumptions from the programme theory and the transcribed interviews, texts, and programme documents. Using Pawson's (2006) questions for complex programmes as a way of organizing the evidence, the different understandings

led to some intended and several unintended outcomes within the realm of the services' direct influence.

For instance, student attendance at the ALLU dropped after the first few weeks of a semester. The causal mechanisms for engaging students appeared to be strengthened when lecturers would endorse the ALLU's noncredit-bearing bridging of academic literacy courses. Conversely, if lecturers allowed their students to use the time allocated to the ALLU as a free period or to catch up on their own syllabuses, this resulted in low student motivation to participate. This helped surface an implicit assumption in the programme theory that would be tested in the evaluation project. The contrasting approaches, which were initially implicit, became particularly evident when the team viewed the programming of the ALLU and the writing center separately.

The writing center ran workshops for faculty lecturers to improve their own research writing. These lecturers were the ones to not only endorse but also to request programming for their students. This major observation regarding contrasting pedagogies and what worked for whom in what contexts and why became evident through *iterative* thinking and an analysis process involving the evaluand. The evidence from the ALLU and the writing center helped explain how these contrasting approaches resulted in different engagement outcomes.

This generative learning process engaged key evaluand members and the project 'client', who had never experienced a learning-focused evaluation before, in contrast to the accountability reporting requirements to funders. These insights and findings would later be used to make organizational changes that should continue to deliver improvements—assuming some basic resourcing conditions.

The community health care case

The EYBS project involved four versions of a theory of change to arrive at a final version that reflected the focus of the evaluation on access to breastfeeding support services at PQCHC by expectant and new mothers. While the initial theory of change was useful for understanding the adherence of mothers to international breastfeeding guidelines, it was too general for the purposes of the evaluation. By working on several refined versions of a theory of change, the evaluators gained a deeper understanding to frame the evaluation with a focus on changes that could be made by PQCHC for improving access to breastfeeding support services. The final version of the theory of change functioned effectively to provide a shared understanding of the focus of the evaluation among the evaluators and client team.

The application of socioecological theory proved beneficial for a mutual understanding among the evaluators and client team about the implementation context for the breastfeeding support services provided by PQCHC. Client familiarity with socioecological theory enabled the evaluators to efficiently develop an effective socioecological model for the evaluation that considered

the expectant and new mothers who use the services; the community of partners, friends, peers, and family members who influence the breastfeeding decisions of mothers; the breastfeeding support services available at PQCHC and within the broader community; and the institutions and policies that can influence the breastfeeding decisions of mothers. The client shared that the socioecological model highlights where PQCHC can have the most impact in supporting access by mothers to breastfeeding support services, and also how PQCHC can influence the other levels for greater effect.

Insights for learning and improvement as part of the evaluation process

Theory-based approaches that involve collaboration between the evaluators and clients from the start to the finish of the evaluation projects support the identification of actional insights to inform decision making by clients during the evaluation process. This contrasts with clients waiting to receive a final evaluation report before implementing beneficial changes based on the evaluation findings. For both cases, the teams found that the highly engaged, iterative online process of theory developing and testing produced insights in *real time* as the studies took place.

The historically disadvantaged institution case

The engagement of different stakeholders in the construction of theory reveals contrasting perspectives that can provide key insights, as explored by the essay from Mulcahy et al. in this volume. In the MUT academic literacies and writing assistance case, it was found that contrasting perspectives (theories) explained findings. The implicit 'remedial' approach in the deficit model thinking resulted in designing large classes for mass learning, but attendance dropped off after the beginning of each semester. So, large classrooms remained empty while students used on-demand, walk-in services for writing support during assignment deadlines. This situation resulted in difficult-to-manage surges and lulls in demand for staff time and other resources, as well as a lack of 'safe space' for student consultations. The dropping number of attendees suggested that the voluntary nature of attendance was insufficient to improve students' access to higher education with success. Indeed, it was the buy-in from lecturers benefitting from writing assistance for advancement of their own careers at the writing center that seeded a response from the lecturers to actively arrange for the students to get regular support from the writing center and the ALLU.

The original implicit program's archetype—centrally located student academic support services providing an advisory or education programme archetype (Funnell & Rogers, 2011)—assumes a self-motivated target audience that receives a service that addresses their stated need, and then encourages others through follow-up. It was the prevailing notion in South Africa that writing was

a skill that should have been acquired before university admittance; this resulted in lecturers viewing students from a *deficit* perspective that needs to be remediated. By contrast, the origins of the newly formed writing center were premised on different philosophical underpinnings that consider all students entering university study to be novice disciplinary members who need to be 'apprenticed' in the discipline and that their lecturers played a key role and needed to be supported to provide this ongoing disciplinary writing 'apprenticeship'.

In the end, the MUT literacy and writing assistance evaluation findings were presented using a cooperative and collective learning approach. Months later, an opportunity arose to present the findings at a conference on higher education with the theme *Internationalisation, Inclusion and Social Justice—Towards a Fairer World.* The key stakeholders in the ALLU and the writing center noted that the evaluation report helped illustrate their respective niche areas. It identified that academic literacies and writing theory have been 'at odds with each other'. ALLU are redesigning their support material in ways that acknowledged the social aspects of learning disciplines such as science, namely 'these issues are the ways of being, behaving, dispositions that students need to develop for ease of access to the disciplinarity of science and chemistry, situated in the social relations' (Madondo et al., 2023).

The identification of these contrasting approaches appeared to be resulting in tensions; however, once the evaluation findings were shared a conversation opened between the ALLU and the writing center that resulted in a clearer role definition and strategy. What emerged is that the ALLU would help introduce pre-tech and first-year students to the literacies they need for the university context with the faculty lecturers engaged to highlight the needs in their disciplines of study. By contrast, the writing center would focus on helping students learn how to write for their disciplines of study. This focuses on helping upper-year students and early career researchers, such as faculty that are still obtaining their PhDs at other universities.[1] Importantly, because the writing center supports the lecturers with their own writing, and the ALLU works on piloting support with certain lecturers, these faculty grew in their understanding of why their students were struggling with their writing. These interactions with the lecturers first, to better understand how to support the students, indicated that the focus on the enabling environment—as identified in the theory of change—was the key to improve engagement.

The expanded and integrated vision of the ALLU and writing center programmes was not present at the beginning of the evaluation study. This

---

1 MUT does not yet have a doctoral program but is pursuing skills upgrading for their faculty so that they can supervise doctoral research and increase their funding from the South African Department of Higher Education.

collaborative theory-based approach appears to have profoundly changed the understanding of how these programmes worked, for whom and why.

The community health care case

The theory of change, including related assumptions, proved to be an invaluable reference for identifying applicable research questions and methods to be applied during the EYBS project. The research methods applied included: a focus group with staff, interviews with staff, a survey with staff, a survey with mothers who recently used the breastfeeding support services, a desktop review of pertinent client documentation, and a literature review of academic and grey literature. The research questions formulated for the evaluation were considered within the context of the theory of change to help ensure alignment with the focus and scope of the evaluation.

Draft interim reports concisely summarizing the key findings from each of the research methods applied during the EYBS project were delivered to the client. This approach enables a client to keep informed about the research findings throughout the evaluation process and enables evidence-informed decision making by a client before the end of the evaluation project. According to the client, the evaluation process supports clients in understanding and appreciating that the learning takes place throughout the evaluation so that the client is not waiting for and putting all emphasis on the final report for the evaluation.

During the EYBS project, the client suggested a sense-making workshop involving the evaluators and client team. The objective of this workshop was for the evaluators to present a summary of the synthesised key findings from all the research conducted during the evaluation. The evaluators welcomed this opportunity to collaboratively review the findings with the client team since this provided a critical opportunity to further contextualize and understand the findings from the research conducted during the evaluation. Also, this provided the client with another opportunity to draw on the research findings for decision making before the delivery of the final report for the evaluation.

As part of project follow-up, the manager of organizational development for PQCHC noted insights related to the programme theory in terms of different stages of access—as well as how the socioecological model helped highlight differences in reach and influence and how different areas could be influenced to achieve greater effect.

Challenges

The conduct of theory-based evaluations in the two cases noted faced significant challenges. Two of the main challenges are considered here:

The communication of a theory online can be difficult since nuances and subtleties can sometimes be lost when people are not dealing in person and onsite. The bandwidth is simply narrower. There is no easy answer to this problem. Both teams addressed this difficulty by taking the time to engage, revise, do homework, revise, engage again, and repeat the process over several months.

Linked to the above, communications are typically less vivid and less frequent online—which can make theories expressed in written form appear to be more fixed than they are in the minds of all concerned. It is important for evaluators and all stakeholders to consider the theories, *not* as 'blueprints' for testing but rather as 'nimble heuristics' to provoke dialogue and to be subject to constant review, reassessment, and refinement.

Conclusion

The two cases show how theory-based approaches applied entirely online using a highly engaging, iterative, and collective learning approach can allow new evaluators to structure evaluation planning, implementation, and reporting that produces important insights and meets client needs. Specifically, online engagement with clients leads to trusting relationships. The flexibility of interacting across a video platform enabled the novice evaluators to meet their client regularly, either across town or across time zones and continents.

The cases illustrate how programme theories can be used to promote stakeholder engagement, as illustrated when using the theory of change along with other heuristic devices like the ecological model. These high levels of engagement facilitate iterations in theorizing and modeling, which in turn creates 'thick descriptions' (see Essay 22 by Davies, Mani, and Hobson in this volume), clarifies thinking, and produces insight. The iterative engagement and process of sharing insights and findings created a space for discussion amongst the educators at MUT and the staff and management of the PQCHC that helped to form insights that were produced as part of the study process. This helped decision makers in real time.

As the post-COVID world reconciles the way it does everything, including evaluation, the early insights from these student practicum cases offer a promising template for further consideration and development.

References

Artemeva, N. (2008). Toward a unified social theory of genre learning. *Journal of Business and Technical Communication*, *22*(2), 160–185.

Baba, J., Emdon, H., McVeigh, M., & Sow, T. (2020). *Academic literacy initiative of the Mangosuthu University of Technology: Tools and preliminary findings*. Major assignment 2, Qualitative Research, April 14, Carleton University.

Boughey, C. (2016). Academic literacy and the decontextualised learner. *Critical Studies in Teaching & Learning*, *4*(2), 1–9.

Centers for Disease Control and Prevention. (2021). *The social-ecological model: A framework for prevention*. Retrieved from www.cdc.gov/violenceprevention/about/social-ecologicalmodel.html

Dahler-Larsen, P. (2018). Theory-based evaluation meets ambiguity: The role of Janus variables. *American Journal of Evaluation*, *39*(1), 6–23.

Funnell, S., & Rogers, P. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. Jossey-Bass.

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, *30*(4), 693–722.

IDEO. (2015). *The field guide to human-centered design: Design kit* (1st. ed). Design Kit.

Madondo, N. E, Khumalo, L. and Mbili, Z. (2023). Legitimation Code Theory as a Lens for Students' Academic Support at Mangosuthu University of Technology. *Proceedings of The Focus Conference (TFC 2022)*. Springer Nature, Advances in Social Science, Education and Humanities Research. DOI 10.2991/978-2-38476-006-0_7

Manzano, A. (2016). The craft of interviewing in realist evaluation. *Evaluation*, *22*(3), 342–360.

Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, *30*(2), 119–142.

Mayne, J. (2017). *The COM-B theory of change model (V3)*. Retrieved from www.researchgate.net/publication/314086441_The_COM-B_Theory_of_Change_Model_V3

Michie, S., van Stralen, M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*, 42.

Pawson, R. (2006). *Evidence-based policy: A realist perspective*. Sage.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). *Realist synthesis: An introduction*. ESRC Research Methods Programme, RMP Methods Paper 2. University of Manchester.

Punton, M., Vogel, I., Leavy, J., Michaelis, C., & Boydell, E. (2020). *Reality bites: Making realist evaluation useful in the real world*. CDI Practice Paper 22, March. Institute of Development Studies. Retrieved from www.ids.ac.uk/publications/reality-bites-making-realist-evaluation-useful-in-the-real-world/

26 Theory of change as a tool for tracking Intensive Family Programme developments in Whitetown

Jane Mulcahy, Catherine Naughton, and Seán Redmond

Introduction

The involvement of children under the age of 18[1] in serious crime, often facilitated by their relationships with adult crime networks, is a significant problem in Ireland. Up to one thousand children are estimated to be embedded, or at risk of involvement, in criminal networks at any one time (Naughton & Redmond, 2020). These children comprise a tiny fraction of those who come into contact with the youth justice system, but they account for a substantial percentage of youth crime. Developing effective interventions to enable such children to desist from crime and other antisocial behaviour is a significant challenge for policy-makers and practitioners.

This essay focuses on the theory of change for the Intensive Family Programme pillar in Whitetown, a real, but anonymised urban community in Ireland. This pilot programme consists of an intervention for children involved in adult crime networks in Ireland, initially funded from 2020 to 2023.

In our essay, we discuss the co-design of the theory of change with staff of the service provider and the need to adapt key aspects of the programme for the local context. We commence by introducing the Whitetown programme and describe stages in the design approach, the sequences involved in developing the theory of change, and the initial testing; this demonstrates the value of the theory of change for charting programmatic changes of tack necessitated by on-the-ground testing. Our account of this process provides an insight into the potential challenges involved in the implementation of dynamic interventions in the context of partial evidence with reference to theoretical and empirical research.

---

1 For the purposes of this essay, an individual under the age of 18 is a 'child', as per section 3 of the *Children Act, 2001*.

Background to the Whitetown programme

Recent Irish research undertaken by the Research Evidence into Policy Programmes and Practice (REPPP)[2] at the University of Limerick investigated child involvement in adult crime networks in a real, but anonymised town referred to as Greentown (Redmond, 2020). Greentown, and two subsequent replication case studies—Redtown (Naughton et al., 2020) and Bluetown (O'Meara Daly et al., 2020)—demonstrated the criminogenic impact of adult criminal networks on children. These networks promote serious criminal activity among children and threaten children's safety, exposing them to exploitation, coercive control, and violent victimisation by network-involved adults.

Designing a solution

A deliberative design process (Rittel & Webber, 1973; Sparrow, 2008) was employed to develop a novel, context-specific intervention involving international academics researching illicit networks, policymakers, state agencies, and community-based organisations.[3] The intervention is based on four complementary but distinct programme pillars: (1) network disruption, (2) pro-social opportunities, (3) community efficacy, and (4) an intensive family support programme (Figure 26.1). The first pillar seeks to reduce the influence of criminal networks on children, the second offers participants opportunities for pro-social activity, and the third seeks to support and build community efficacy. The fourth pillar aims to improve family functioning for the children most embedded or at risk of involvement in serious network-related crime, utilising functional family therapy (FFT), an established, evidence-based, 'clinically creative' programme (Sexton, 2016), described as a 'best practice' treatment of choice for working with adolescents and families of externalising disordered adolescents (Sexton & Alexander, 2004, p. 27).

These 'mutually supporting pillars' (REPPP, 2017, p. 5) were identified as most likely to meaningfully respond to the problems presented by criminal networks in Ireland.[4] The pillars and underlying assumptions were scrutinised by simulation of real-life implementation scenarios.

The design process included rigorous stress testing by police, probation, and child protection social workers. On-the-ground innovation was anticipated at the design stage. The overarching programme is intentionally *part-developed*

---

2 See https://ulsites.ul.ie/law/node/106891 for REPPP publications and media coverage.

3 Almost 90 experts were involved in the consultative process, contributing approximately 590 hours of input.

4 This pillar was originally called Child Agency. It was renamed to reflect the fact that family dynamics have an impact on child self-determination, as well as their well-being, functioning, relationality, offending behaviour, and vulnerability to network involvement.
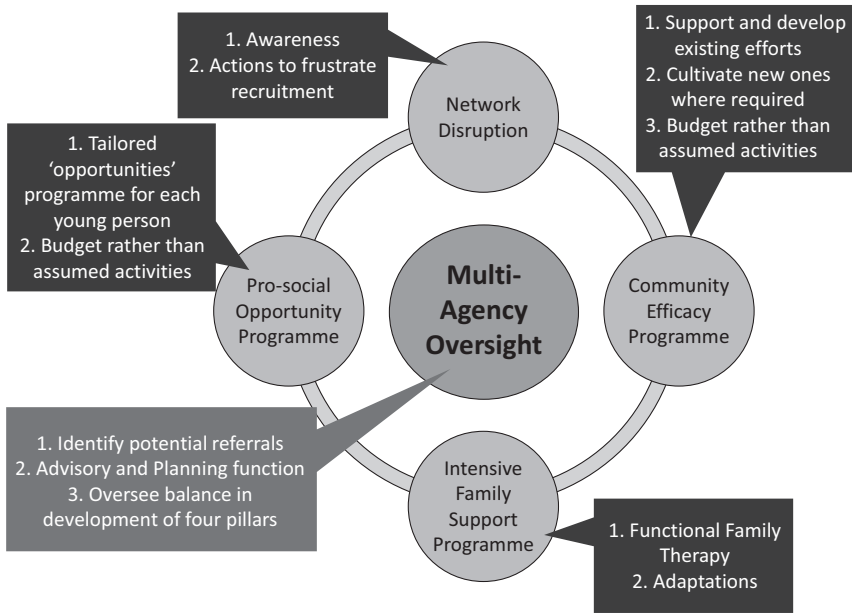
Figure 26.1 Whitetown programme pillars as finalised at procurement stage

to allow for adaptation during implementation to fit the local context. The part-developed approach seeks to retain the core integrity of the programme while enabling service providers, local residents, and community workers in the pilot sites to actively participate in the iterative design process.

Consultees noted the limitations of evidence-based programmes that have been subject to rigorous evaluation methods, such as randomised control trials, particularly for dealing with complex problems in complex settings. While programmes such as FFT[5] are important for idea generation, experts cautioned against taking an 'off the shelf' programme and dropping it into unique community contexts (REPPP, 2017). Adaptations and evidence-informed local crafting would be necessary.

Phase 1 of establishing the overarching model for the intervention involved the development of the four interdependent pillars. The programme aims to: (1) frustrate the recruitment of children into criminal activity by network-based adults and reduce the effects of those networks on their involvement in crime,

5 See www.fftllc.com/about-fft-training/clinical-model.html

and (2) provide meaningful and practical routes out of crime for children already embedded in a criminal network.

Co-designing the Intensive Family Programme theory of change

A theory of change approach was adopted to aid our understanding of the Greentown programme's *contribution* in several complex systems (Mayne, 2008), helping to generate ideas as a group (Zhou et al., 2019) and surface the underlying rationale(s) to support programme 'planning, implementation, and assessment' (Reinholz & Andrews, 2020, p. 1). We envisaged that our two programme aims would be achieved 'through a logical sequence of intermediate outcomes' (Maini et al., 2018). While experimental design is useful for determining whether a programme has impact by strengthening causal inferences, theory-based evaluations that test a robust theory of change may be more effective at unpacking with greater precision how, why, with whom, and in what circumstances a programme works. A theory of change describes the 'causal assumptions behind the links in the pathway—what has to happen for the causal linkages to be realised' (Mayne, 2019, p. 172).

The theory of change for the Intensive Family Programme was co-designed using the online platform *Microsoft Teams* during the COVID-19 pandemic. Participants at the theory of change development sessions included three members of the research team and three practitioners from the Intensive Family Programme provider. A total of four virtual sessions between February and April 2021 resulted in the co-design of an initial draft theory of change using Miro.[6] A review was held in November 2021 to track the implementation of the theory of change on the ground and determine whether tweaks were required to reflect the need to pivot in the light of context-specific realities for children and families involved in the programme.

The first signs of discrepancy between FFT aims and Whitetown programme aims arose at the initial theory of change session. A request that therapists attempt to identify provisional, proximal outcomes led to service-provider apprehension on the basis that focusing attention on variables being measured might distract the therapists and thereby reduce fidelity to FFT. Researchers explained that efforts applied to collecting such data could assist with timely identification of crucial soft outcomes not captured in routine monitoring by the evidence-based programme. These include executive function (Cheng et al., 2019) and agency (Maruna, 2001), upon which hard criminal justice outcomes rest (Andrews & Bonta, 1998).

At subsequent sessions, Intensive Family Programme staff observed that positive behavioural change and openness to pro-social opportunities would depend

---

6 See www.miro.com.

firstly on engagement, followed by a sense of safety and trust in the presence of the FFT therapist, who is deemed to be 'the expert on the change process' (Sexton & Alexander, 2004, p. 47), and better relational dynamics between the caregiver and the young person and improved emotional self-regulation. They also identified examples of skills that adolescents and families were likely to obtain from programme participation, such as negotiation skills, improved interactions with authority figures around negotiating boundaries, anger management, risk acknowledgment, and increasing risk aversion.

The research team presented the draft theory of change to the multiagency Whitetown Local Advisory Committee (which provides a governance function for the project) for feedback, explaining the various stages, iterations, additions, and sequencing changes. The committee observed that engagement would be a necessary prerequisite. It was also agreed that the theory of change would need to be kept under review, informed iteratively by practice on the ground with live cases.

The theory of change development highlighted programme discrepancy, prompted changes of tack, and redeployment of human resources

Since the theory of change was developed, the Whitetown Intensive Family Programme manager has facilitated six separate case reviews with therapists who provide a progress update on the adolescents and families with whom they are working. The transcribed recordings are in the process of being analysed for consistency with the draft theory of change. The next step will be to refine and operationalise the theory of change to determine appropriate measures for evaluation.

In reviews, therapists repeatedly noted that the safety of child participants and their family members is a precondition for positive behavioural change. While FFT was designed to work with 'treatment-resistant' individuals, the model says little about how a lack of safety might impede engagement and motivation to engage with therapy (Sexton & Alexander, 2004, p. 3). Initial surface analysis of the case reviews suggests that a lack of a 'felt sense' (Levine, 1997, p. 69) of safety (Porges, 2009) may pose a barrier to therapy readiness.

Much of the therapists' time has been devoted in the first few months to practical case management—that is, trying to meet the basic physiological and safety needs of participants, whether this means concrete assistance to find new school placements, providing small amounts of money to buy food for dinner, helping with housing problems (for example, a missing front door in one case), or other survival needs. In a small number of cases, family stress is further compounded by external threats due to the young person accruing a drug debt.

The theory of change development provided a safe space to test assumptions and re-engineer the programme in real time, particularly around the engagement

process. Attempted therapeutic or behavioural change with children and caregivers seems to be especially challenging when trauma and danger within the home is ongoing (for example, due to domestic violence). This learning may mean that an additional pre-FFT phase is required in the Whitetown context; some trauma therapists refer to this as *safety and stabilisation* (Fisher, 1999; Lohrasbe & Ogden, 2017).

Theory of change discussions disclosed a need that may not be fully served by orthodox FFT regarding so-called 'hard to reach' children (Department of Justice and Equality, 2021, p. 12). Participants deliberated whether an outreach component might be necessary to connect with this cohort (Mulcahy, 2021). Families with children who are embedded in networks might be unwilling to avail of the offer of therapy or a mentor, perhaps due to drug-related intimidation (DRIVE, 2021) or direct family involvement in network activity (Redmond, 2020). This real-life demand articulated through the theory of change process presented potential conflicts with the fidelity requirements of FFT, which envisages a short engagement and motivation phase, usually between three and six sessions (Sexton & Alexander, 2004, p. 102), and a therapeutic focus on family relating patterns and competency development rather than fostering dependency (Sexton, 2016; Sexton & Alexander, 2004).

The likely necessity for on-the-ground innovation was anticipated at the deliberative design stage. The four-pillar programme was part-boiled to allow for adaptation during implementation. Moreover, evidence-informed changes of tack were enabled by the theory of change development that rendered programme assumptions explicit and exposed the need to pivot in light of the complex presentations of teenagers and families who voluntarily opted into the programme.

While the voices of former network-involved individuals were missing from the expert-led deliberative design process, REPPP has begun the process of harnessing perspectives of 'lived experience professionals' (Brierley, 2021, p. 109) that may, in due course, lead to programme modifications, where necessary and appropriate. We have convened an expert advisory workshop with a diverse membership to discuss real-world problems experienced by the Whitetown pilot project. There are several people in this group with histories of adverse childhood experiences (Felitti et al., 1998), adverse community environments (Ellis & Dietz, 2017), addiction, offending behaviour, and imprisonment who have successfully seized 'hooks for change' (Giordano et al., 2002) and desisted from crime (Maruna, 2001). These individuals are likely to have pertinent insights about barriers to engagement and therapy readiness, which may lead to revisions of the theory of change.

Conclusion

This essay has described the Whitetown programme, the expert-led design process, and learnings from the Intensive Family Programme theory of change

development. We gave an account of how the process of co-designing a theory of change provided early indications of discrepancy between Greentown needs and FFT fidelity requirements, and the potentially limited capacity of FFT to cater for an enhanced engagement phase for unsafe young people who are slow to trust.

We explained that the Greentown programme designs were *emerging* as the process developed beyond the point of commissioning the service in Whitetown. Indeed, work on the ground by therapists endeavouring to apply FFT with fidelity quickly revealed the need to prolong the engagement stage to undertake more intensive therapeutic case management. This adjustment was necessary due to observed trauma-related safety issues that appeared to hamper responsiveness to therapy in line with FFT's projected timeline.

To date, the therapists have primarily worked with children on the periphery of the network. Further adaptations to the Intensive Family Programme pillar are likely as therapists work progressively with children who are more deeply embedded in the network, since they might feel greater loyalty to network associates than to their own families or could be more at risk of external safety threats that may hinder engagement with the programme. The development of the theory of change provided transparency by documenting assumptions and proposed processes. Cognisant that our body of evidence is partial and incomplete, subsequent versions of the theory of change informed by the realities of service delivery in Whitetown's complex environment and relevant lived experience insights from expert workshops will be captured and carefully documented for evaluation purposes.

References

Andrews, D., & Bonta, J. (1998). *The psychology of criminal conduct* (2nd ed.). Andersen Publishing.

Brierley, A. (2021). *Connecting with young people in trouble: Risk, relationships and lived experience*. Walterside Press.

Cheng, J., O'Connell, M., & Wormith, S. (2019). Bridging neuropsychology and forensic psychology: Executive function overlaps with the central eight risk and need factors. *International Journal of Offender Therapy and Comparative Criminology*, *63*(4), 523–542.

Department of Justice and Equality. (2021). *Youth justice strategy, 2021–2027*. Department of Justice and Equality.

DRIVE. (2021). A data-driven intervention model to respond effectively to drug-related intimidation and violence in communities in Ireland. Retrieved from http://driveproject.ie/wp-content/uploads/2023/01/DRIVE-Report.pdf

Ellis, W. R., & Dietz, W. H. (2017). A new framework for addressing adverse childhood and community experiences: The building community resilience model. *Academic Pediatrics*, *17*(7), S86–S93.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household

dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American Journal of Preventive Medicine*, *14*(4), 245–258.

Fisher, J. (1999). *The work of stabilization in trauma treatment*. The Trauma Centre Lecture Series.

Giordano, P., Cernkovich, S., & Rudolph, J. (2002). Gender, crime, and desistance: Toward a theory of cognitive transformation. *American Journal of Sociology*, *107*(4), 990–1064.

Levine, P. (with Frederick, A.) (1997). *Waking the tiger: Healing trauma*. North Atlantic Books.

Lohrasbe, R. S., & Ogden, P. (2017). Somatic resources: Sensorimotor psychotherapy approach to stabilising arousal in child and family treatment. *Australian & New Zealand Journal of Family Therapy*, *38*(4), 573–581.

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 Global Health*, *3*(1). https://doi.org/10.1136/bmjgh-2017-000617

Maruna, S. (2001). *Making good: How ex-convicts reform and rebuild their lives*. American Psychological Association.

Mayne, J. (2008). *Contribution analysis: An approach to exploring cause and effect.* ILAC Brief 16.

Mayne, J. (2019). Revisiting contribution analysis. *Canadian Journal of Program Evaluation*, *34*(2), 171–191.

Mulcahy, J. (2021). Relentless caring: Trying something new: An evaluation of the Targeted Response with Youth TRY Project. Retrieved from www.drugsandalcohol.ie/34556/1/TRY-Evaluation.pdf

Naughton, C., & Redmond, S. (2020). National prevalence study: Do the findings from the Greentown study of children's involvement in criminal networks (2015) extend beyond Greentown? Retrieved from https://researchrepository.ul.ie/articles/report/National_prevalence_study_do_the_findings_from_the_Greentown_study_of_children_s_involvement_in_a_criminal_network_extend_beyond_Greentown_/19830238

Naughton, C., Redmond, S., & O'Meara Daly, E. (2020). Lifting the lid on Redtown: A replication case study, which investigates the contribution of engagement in a local criminal network to young people's more serious and persistent offending patterns. Retrieved from https://researchrepository.ul.ie/articles/report/Lifting_the_lid_on_Redtown_a_replication_case_study_which_investigates_the_contribution_of_engagement_in_a_local_criminal_network_to_young_people_s_more_serious_and_persistent_offending_patterns/19830250

O'Meara Daly, E., Redmond, S., & Naughton, C. (2020). Lifting the lid on Bluetown: A replication case study, which investigates the contribution of engagement in a local criminal network to young people's more serious and persistent offending patterns. Retrieved from www.ul.ie/research/publications/lifting-the-lid-on-bluetown-a-replication-case-study-which-investigates-the

Porges, S. (2009). The polyvagal theory: New insights into adaptive reactions of the autonomic nervous system. *Cleveland Clinic Journal of Medicine*, *76*, S86–S90.

Redmond, S. (2020). Lifting the lid on Greentown Version 2—Why we should be concerned about the influence criminal networks have on children's offending

behaviour in Ireland. Retrieved from https://researchrepository.ul.ie/articles/report/Lifting_the_lid_on_Greentown_Version_2_Why_we_should_be_concerned_about_the_influence_criminal_networks_have_on_children_s_offending_behaviour_in_Ireland/19830277

Reinholz, D. L., & Andrews, T. C. (2020). Change theory and theory of change: What's the difference anyway? *International Journal of STEM Education*, *7*, 2.

REPPP. (2017). Greentown programme design. Workshop 3 Report, December 12–13. Available from the author on request.

Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169.

Sexton, T. (2016). Functional family therapy: Evidence-based and clinically creative. In T. L. Sexton & J. Lebow (Eds.), *Handbook of family therapy* (pp. 250–270). Routledge/Taylor & Francis Group.

Sexton, T., & Alexander, J. (2004). *Functional family therapy clinical training manual*. Strategies.

Sparrow, M. (2008). *The character of harms: Operational challenges in control*. Cambridge University Press.

Zhou, X., Zhai, H. K., Delidabieke, B., Zeng, H., Cui, Y. X., & Cao, X. J. (2019). Exposure to ideas, evaluation apprehension, and incubation intervals in collaborative idea generation. *Frontiers in Psychol*ogy, *10*, 1459. https://doi.org/10.3389/fpsyg.2019.01459

Part 7

Conclusions around theories of change in reality

Andrew Koleros, Marie-Hélène Adrien, and Tony Tyrrell

This book sets out the responses from a group of international practitioners and experts to three questions posed by the editors around how they are using theories of change in their work, what they have found to be the limits of their use, and what further adaptations they feel are needed so that theories of change remain relevant and useful in the future. The responses we received are, given the diversity of the contributors, understandably diverse with reference to perspective, emphasis, and the context within which theories of change are used. Broadly, their perspectives indicate that theories of change have long represented a valuable approach in both programme planning as well as in evaluation design and practice; however, the authors also make clear that theories of change are often misunderstood and misused. In particular, the authors draw attention to the use of theories of change as a static product rather than an iterative, inclusive process that recognises the centrality of multiple stakeholder perspectives in evaluation design and practice. The authors also highlight the challenges involved in the application of theories of change in contexts defined by heightened complexity.

These insights tell a compelling story about how the uses of theories of change have shifted over time, particularly over the last 10–15 years, along two important and overlapping dimensions:

- The relative understanding and emphasis of a theory of change as an approach to infusing causal logic and evaluative thinking into a programme or evaluation design process answering the question 'how change happens' (theory of change as 'process') versus a diagrammatic or pictorial model of 'did the programme work' and how a particular intervention intends to bring about change (theory of change as a 'product').
- The situations for using a theory of change from more simple, bounded programme contexts to more open and complex systems or contexts.

As these two dimensions have shifted, so too has the nature by which different types of stakeholders engage with theories of change, particularly individuals who are most impacted by the problems that interventions are intending

to address. Reflecting this, we note in this concluding chapter how theories of change can be used as dynamic tools that allow evaluators to meaningfully involve stakeholders, although this is not always done in practice, and we discuss how evaluators and programme and evaluation commissioners might better maximise their potential to navigate complexity.

Process versus product

A theory of change approach can offer a flexible and adaptable process that embeds evaluative thinking and surfaces perspectives from a diverse range of programme stakeholders.

The authors make it clear that theories of change are being used for multiple applications (planning, implementation, communication, and evaluation) in multiple contexts (international development, public policy, private sector, and social enterprise). Across these diverse applications and contexts, a clear theme emerged: the process used to develop theories of change is just as valuable—and in some contexts can be more valuable—than the product that comes out of that process. While it was widely recognised that a good theory of change makes a programme evaluable, authors also described theory of change development as a powerful reflective process that makes implicit logic explicit. Leeuw, for example, has equated this process as a dance between mechanisms and context, where the focus is not on identifying the expected results of an intervention but rather to understand how change happens.

Authors in this book described the power of this process when using theories of change for designing programmes (Koleros, Tangelder), programme management and implementation (Oldenbeuving, Freer), or evaluative purposes (Perrin, Andersson, Hawkins, & Tolmer). When done well, at the programme design phase the process ensures that a programme is underpinned by sound social science theory; that programme teams sufficiently appreciate the context in which the programme is being implemented; and that assumptions about how, why, and for whom a programme is intended to bring about change are made explicit (Lemire, Van der Knaap). Using theories of change as the basis for evaluation can then support evaluators in examining the logic and rationale underpinning a programme and test it against alternative hypotheses (Levine, Delahais, Palenberg).

Inclusive stakeholder engagement can enrich the process and enhance the quality of any product (Lemire). Authors described the benefits of engaging stakeholders in creating or validating the theories of change underpinning any intervention being evaluated, whether a project (Mulcahy et al.), a programme (Montague et al.), or multiple competing courses of action (Davies et al.). Such engagement may also provide benefits to those participating in the exercise; for example, Olejniczak and Lyubashenko's work, although experimental, suggests

that conscious, structured stakeholder engagement has the potential both for enhanced empowerment and joint ownership of the evaluation process as well as ensuring that the final product reflects the real needs and priorities of all concerned. Mulcahy et al. observed positive behavioural change amongst those engaged in a theory of change process, citing examples of skills that adolescents and families were likely to obtain from engaging in the process, such as negotiation skills and improved interactions with authority figures.

> **As theories of change have become more widely adopted in evaluation and other practice (for example, planning and design), there has been a shift in emphasis from generating value through the 'process' towards delivery of a stand-alone 'product'. This often limits its usefulness and, in some cases, has led to misuse.**

Authors identified relatively innocuous misuse of theories of change in practice, such that they become a tick-box exercise or a donor-mandated deliverable tied to funding authorisations (Freer) and disassociated from the programme design or conceptualisation process. However, in many cases these misuses can cause harm to a programme, whether that be by confusing a theory of change with a logframe leading to unrealistic expectations about the results that can be achieved in a given time horizon or by using the theory of change as a static model that does not accommodate uncertainty when implementation context requires adaptation. Perrin decries practice that he has observed wherein the theory of change is used as 'a rigid input-output model, with indicators and targets identified in advance' rendering it a glorified logframe. It is an interesting comparison to make, as we recognise that the logframe is a useful tool that, through its own overuse, has become 'institutionalised' and in some ways suffered a fate like that being discussed in this book around theories of change. A wider observation one may make is that perhaps any widely used approach or model is likely to be subject to nonuse or misuse unless this is guarded against—a topic we will return to at the end of this chapter.

Returning to the topic at hand, authors attributed some of this misuse and abuse to an overemphasis on the 'product' aspect of theories of change. Indeed, if you ask many practitioners to describe what a theory of change is, their response will mostly include some description of a concrete product, such as 'a bunch of boxes and arrows' or 'a landscape diagram in MS PowerPoint or MS Word'. This framing de-emphasises the value of process and engagement and places the focus more on how the theory of change is visually depicted or represented. While visualising a theory of change is important and merits attention, the depiction of the final product should not come before or at the expense of the process to develop it. A number of the authors in this book have provided guidance and pointers on how to honour the process, including, for example, through stakeholder engagement and dialogue (Rogers, Dahler-Larsen). Ling's

sign-off to his essay posits a set of principles for maximising the benefits of and minimising the risks for evaluators in developing theories of change, as follows: engage stakeholders, incorporate systems thinking, iterate the theory of change, prioritise theory building, and be humble and curious.

Similarly, attempting to develop a theory of change to conform to a predetermined final format or template misses the point of the approach entirely and all but ensures that the maximum value of a theory of change cannot be generated. One cannot shoehorn the results of a theory of change process into a predetermined format or template; attempts to do so put unnecessary constraints on the process that result in some of the uses and misuses that authors described in their essays.

Further, overemphasising the theory of change product more than the development process does not make good sense, as there are multiple ways to depict a theory of change and any one intervention may have multiple depictions of its theory of change depending on its uses, whether that be for design, evaluation, or communication purposes to external audiences (Koleros). For instance, Morkel, Lima & Lafer, and Dahler-Larsen describe the 'rote' use of linear models of theories of change—typically framed in one-dimensional graphic form and used mostly for communication purposes—as insufficient to address the complexity and magnitude of development challenges.

As these authors highlight, a recognition of the need for multiple ways to represent and depict theories of change becomes even more important when using theories of change to widen the lens on our understanding of how change happens from closed and bounded programme contexts to more open systems, where boundaries are more subjective and contested, and our understanding of underlying system dynamics are less certain and predictable. This is further explored in the following section.

Simple versus complex

This shift in emphasis from 'process' to 'product' has led to limitations in the usefulness of the approach, particularly as practitioners use theories of change to understand how change happens in situations that have moved from simple to more complex dynamics and behaviours.

Over time, evaluators' work has progressively moved from evaluating single interventions where cause-and-effect relationships are relatively simple and predictable to work in more uncertain and less predictable situations, particularly in international development but also in national planning and related areas (Lima & Lafer, Morkel, Rogers). Evaluation commissioners and other governance systems increasingly require evaluators to evaluate complex interventions aimed

to address complex, interconnected global challenges related to climate change, insecure food systems, the spread of infectious diseases, and growing economic inequalities, among others (Dahler-Larson, van der Knaap).

While the nature and types of programmes and interventions that evaluators are commissioned to evaluate have evolved, the use of theories of change in these new situations has not evolved at the same pace. As practitioners have responded to these new situations, they have confronted challenges in trying to apply tools and approaches that were designed and intended for bounded programme contexts to more complex systems without sufficient adaptation of the process or setting the context for use appropriately (Andersson). Several authors discussed that when applied beyond the project or simple programme 'box' to situations characterised by complex system dynamics, the use of theories of change to understand and manage complexity is challenged (Morkel, Lima & Lafer, van der Knaap, Houlberg & Rieper).

In many cases, it appears that the evaluators' solution to adapting theories of change is more consistent with complicated rather than complex situations (Rogers, 2008). In practice, this means that evaluators use theory of change products designed for more simple situations and attempt to layer on more and more interconnected pathways, steps in pathways, feedback loops within and across pathways, connections with other interconnected systems, and so on, as a way to 'deal with the complexity' they face. This results in certainly more complicated theories of change that capture and depict more things within an overall theory of change product, but often do not provide sufficient insights into how and why change happens within a particular situation, thus reducing the usefulness of the overall theory of change approach.

Attempting to capture complicated system dynamics in a theory of change product is different from using a theory of change to guide a process of exploring our understanding of how change happens in a complex system and how we might usefully intervene to address underlying system dynamics. In these situations, the limited utility of the theories of change can contribute to some of the misuses and harms described previously. As a consequence, some authors in this edition concluded that when used as a product over an approach, theories of change may have reached their limits in terms of their ability to embrace and explain complexity (Dahler-Larson, Palenberg).

Others argue, however, that these challenges are more symptomatic of the overreliance on the 'product' versus the 'process' aspect of theories of change, without sufficient attention put on how to adapt *the process* of developing or reconstructing theories of change for situations characterised by complex system dynamics. As described above, that process should result in theories of change that are underpinned by sound social science theory of how change happens in complex adaptive systems, not just more complicated diagrams.

Many authors acknowledged that the conventional ways of visually depicting theories of change designed for more simple situations are not appropriate

for more complex challenges and wider societal processes. However, rather than concluding that this means theories of change have reached their limits, they argue that, for these endeavours, evaluators need to develop and learn an entirely new set of techniques consistent with developing theories of change that embrace complexity and new ways of visually depicting causal logic consistent with 'how change happens' in more complex systems. Indeed, this book provides insights into what these processes might look like (Rogers), as well as practical examples of how these can be applied in reality (Davies et al., Koleros, Leeuw). However, these new approaches and techniques are better classified as 'emerging practice' as opposed to 'business as usual'. To get there, we need to shift not only how evaluators adapt and use theories of change for complex interventions tackling complex problems, but also how evaluation commissioners understand and appreciate the usefulness and benefits of these emergent approaches. This point is further elaborated in the 'useful tips' listed by Gates et al. in the Foreword of this book.

> **Reflecting on this current state of practice, authors suggest the need to reappraise, reimagine, revisit, and regenerate the uses of theories of change as a process over product, particularly in systems where complexity is high and where stakeholder engagement is a critical component of the process.**

Notwithstanding these critiques, a core conclusion is that theories of change have the potential to remain a relevant and potentially cutting-edge tool; however, to meet the challenge of increased complexity in their application (and to avoid debasement of their value through misuse), there is a need to adapt the use of theories of change through enhanced socialisation, stakeholder engagement, and democratisation.

Reflecting Montague, Emdon, and Grabinski's findings that process was the fundamental benefit of a theory-based approach, many authors in this book advocate the need to reassert 'process' in the use of theories of change in order to fully embrace its learning orientation. Tangelder urges a reimagining of how the tool is used to, in part, drive learning and adaptation. And that requires 'intentional design and commitment to deep learning' on the part of the practitioner to guide the choices as well as the type and level of engagement with stakeholders throughout the process of developing a theory of change. Gaarder and Oldenbeuving urge for a regular revisiting of the theory of change by teams so that it can be used for learning and adaptation.

This process orientation is even more critical for the increasing use of theories of change for understanding how change happens in situations characterised by high levels of complexity. The authors that have contributed to this book have opined that meeting the challenge of complexity will require a more contextualised, socialised, and indigenised approach to the development of the theories

of change underlying related initiatives. For example, when viewing a situation systemically—whether from an intervention or evaluation perspective—it is necessary to consider three key elements: relationships, perspectives, and boundaries (Williams & Hummelbrunner, 2010). From a theory of change perspective, this means who participates in the process matters, as different perspectives will surface different types of interrelationships and boundaries (Davies et al., Lemire). Wide stakeholder participation in a theory of change process collectively allows the group to explore jointly the change pathways leading from the resources allocated for an initiative to the expected outcomes and impacts. Ideally, such approaches should be broadly inclusive of stakeholders, including intended beneficiaries of planning and associated interventions (Morkel, Lima & Lafer, Olejniczak & Lyubahenko, van der Knaap, Houlberg & Rieper, Dahler-Larsen, Levine, Delahais, Ling). In practice, given the messiness of reality, as well as other factors such as time and resource constraints, it may not always be possible to be entirely inclusive—but the ambition and desire to be as inclusive as possible should be nonnegotiable.

Given the application of theories of change in increasingly complex settings, their ongoing value will depend not only on the logic of causal analysis but also on how the evaluative process is socially constructed, taking into account political and organisational factors and what is referred to as a more dialogue-oriented approach (Dahler-Larsen, Ling, Levine, Delahais). This more inclusive approach will, they argue, contextually sensitise and enrich the theory, and it can build a sense of ownership and consensus that, in turn, can facilitate fidelity-observed implementation. Olejniczak & Lyubahenko present a mechanism through which citizens can be involved in inputting into and creating the theory of change, a process that can enhance buy-in, help generate mechanisms to ensure implementation fidelity, and so on. All these aspects—process driven, dialogue-oriented, and stakeholder inclusion—are particularly important when the context is fluid, rapidly changing, complex, or conflict affected (Hassnain). In that type of scenario, the relationship between causes and effects is so complex that plans and strategies are likely to require regular reappraisal with associated adaptation of relevant theories of change.

Behaviour change and theory of change practitioners

Within the world of behavioural science, the behaviour change wheel or COM-B framework (Michie et al., 2011) posits that effective behaviour change interventions should address three components—capability, opportunity, and motivation. Capability refers to the individual's physical and psychological capacity to engage in the behaviour. Opportunity refers to the environmental factors that facilitate or hinder the behaviour. Motivation refers to the individual's internal drive to engage in the behaviour. Drawing on this framework, we recognised that shifting theory of change practice among the practitioner

community—evaluators and evaluation commissioners—will require changes in their capabilities, opportunities, and motivations. We offer some suggestions on how this might be achieved in the remainder of this chapter.

Improving the technical knowledge and facilitative skills of theory of change practitioners

Several authors felt that there is a lack of capacity-building activities around theories of change that fully describe how they can be used as an approach for multiple purposes: planning and design, evaluation, and learning (Williams, Freer, Perrin). Capacity development efforts should focus on how to develop and describe theories of change, taking into account the terminology and how we talk about the approach—for example, using a term such as *storytelling* rather than *theory* where the latter may appear difficult or 'threatening' to certain stakeholders—and the specific skills that evaluators would need to do this.

The capabilities required of a practitioner to fully embrace the process aspects of theories of change will increasingly need to move beyond the technical into an embrace of greater levels of stakeholder engagement. These are needed to better understand context and complexity and to draw on multiple perspectives to better frame evaluation efforts. Expert facilitation can ensure that all voices are heard, can identify when appropriate consensus can be reached, and can manage group dynamics so that the group progresses at a pace that will not slow the process. Facilitators need to have empathy and skills in communication and conflict resolution that may or may not be part of current evaluator competencies. They also need the ability to manage the power dynamics that are invariably present between various groups, such as the dynamics between donors and recipients of programme funding, managers and implementers of a programme, marginalised and nonmarginalised groups, or between genders.

Beyond skills development, there is also a need to increase the knowledge base, through more guidance for theory of change development, representation, and use beyond its conventional use as a tool for programme interventions. These need to look different from the existing guidance and could include the use of theories of change as part of wider public-sector planning processes, including in the developing world (Morkel, Lima & Lafer). Specifically, to meet the challenge of complexity new guidance is necessary to help foster a more nuanced, systems-based approach to the use of theories of change (Morkel, Van der Knaap, Houlberg & Rieper, Ling, Palenberg, Delahais).

Creating the opportunities for practitioners to put knowledge and skills into practice

Authors noted that this shift will also require time and resources to bring together individuals to engage in the construction of the theory of change and to come to a common vision. While terms of reference for most evaluations ask the

evaluation team to review or reconstruct the theory of change as a first step, the funds or time required for such an activity, let alone for also involving a range of stakeholders, are rarely allocated or sufficient. And, as noted by Hassnain, in countries with limited or no technology, such engagement requires face-to-face meetings and thus even greater costs.

There is also a risk that those participating in the process may perceive stakeholder engagement as simply symbolic (lip service) if they feel that their voices are not heard or taken into consideration, and this can lead to mistrust between the different parties. Creating opportunities might also mean addressing challenges associated with the bureaucratisation of the evaluation function reflected, for example, in the conflation of theories of change with the logframe, the lack of an outcome orientation consequent on overly simplified models and measurement, as well as other imposed limitations that are comprehensively covered in Perrin and Tyrrell's (2021) volume in this series.

Changing motivations around the use and usefulness of theories of change

All of this requires movement in the motivation of theory of change practitioners and commissioners if they are to successfully navigate complex systems and facilitate the production of theories of change that are 'meaning full' from the perspective of multiple and diversely positioned stakeholders. Contributors to this book have highlighted evaluation practices that are less than helpful in that regard, drawing attention to practice that fails to assert a more sophisticated engagement with complexity and a failure by some to use and further develop thinking and tools—such as contribution analysis—that show promise.

Likewise, the broadest, most inclusive engagement of stakeholders is not simply a question of skills and competencies, such as communication and empathy, but also one of a mindset that accepts that all stakeholders, not simply the most powerful groups, have something legitimate to say. Yes, greater levels of stakeholder engagement can require greater amounts of time and resources. But stakeholder engagement is an imperative rather than a luxury if the practitioner community is to form legitimately constructed judgements about the amount, value, or quality of interventions.

As such, it is incumbent on practitioners to ensure they are adequately equipped to engage with complexity, are trained in emerging approaches, and embrace cross-disciplinary knowledge and learning. It is also necessary for evaluators to inform commissioners and clients of the importance of the broadest range of stakeholder input with a view to rendering theories of change more attuned to complexity, more useful as a tool through which rich feedback can be absorbed and analysed, and leading—iteratively—to adaptation that will render policy and programming more effective. How evaluators and their clients work through these processes is a subject that merits further research. What is clear from the weight of knowledge and experience detailed in the essays presented in

this book is the need for more and richer dialogue to support greater appreciation of when and why theories of change might be of most use and how they can be best designed; that is, to be context-specific, testable, revisable, inclusive/participative, and so on—within limited time and resources.

References

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, *6*, 42.

Perrin, B., & Tyrrell, T. (Eds.). (2021). *Changing bureaucracies: Adapting to uncertainty, and how evaluation can help.* Routledge.

Rogers, P. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29.

Williams, B., & Hummelbrunner, R. (2010). *Systems concepts in action: A practitioner's toolkit*. Stanford University Press.

Index