



**The Feasibility of a Systems-Based Approach in Non-Profit Strategic Planning,
Monitoring and Evaluation: Development and Assessment of the SAMEL Toolkit**

Jessica Horler

HRLJES001

Supervised by Associate Professor Sarah Chapman and Dr Carren Duffy

A thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

School of Management Studies

University of Cape Town

June 2024

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration

I hereby declare that this thesis is my own work, from conceptualisation to execution. To the best of my knowledge, this thesis contains no material written by another person, except where due acknowledgement has been made. I have not previously submitted this thesis in its entirety, or in part, to this university or to any other university.

Signed by candidate

26 June 2024

Student signature

Date

Dedication

To my beloved grandmother Mavis, who always wanted a doctor in the family.

Acknowledgements

To my supervisors, Carren and Sarah, thank you for your consistent support over so many years. Most of all, thank you for your trust and patience. I am incredibly grateful for the freedom you gave me to explore, create, and be exhaustingly curious. Your guidance and advice over the years made me a better researcher and a stronger practitioner. It has truly been a privilege to work with you and learn from you both.

To my family, who never viewed this achievement as anything other than inevitable, thank you for an unwavering and almost pathological belief in my ability to get this done. Thank you for being there for me, for showing up, and for pushing through the worst of times. Mom, you brought me up to believe that an education never ends and is always worth pursuing, no matter how long it takes or how arduous the journey is.

To Cam, this wouldn't have been possible without you. You kept me alive during fieldwork, but, more importantly, you taught me grit and determination. You turned my life inside out so many times and in so many ways, and I'm grateful for all of it. You helped me to develop the life skills that I needed to do this work, and to be a better version of myself overall. You will always be my role model.

To my dear friends Lee, Judy, Jenna, and Amy, thank you for giving me endless support during times that it felt like oxygen. You provided words of wisdom, shoulders and couches to cry on, and an endless supply of motivation. You backed me from day zero, and it has meant the world.

Brett, you were the shining light at the end. Thank you for providing me (and my dopamine-depleted brain) with so much happiness during what is supposed to be a gruelling and miserable time. There aren't enough words to describe what you have done for me. It is a privilege to know you.

To the incredible practitioners who took part in the study, evaluators and NPO teams, I am not allowed to name you, but this research is yours. Thank you for your willingness to join me on this journey, for your openness and your warmth. My fieldwork was such an enjoyable and enriching experience, simply because of you. It was a pleasure and a privilege to work with and learn from all of you.

To the three experts who provided their feedback on the Toolkit, I admire all of you and your contributions to the field of evaluation. To be able to share a part of this thesis with you was an honour.

To the giants who came before me and whose shoulders I stand on, I thank you for paving the way. I am particularly grateful for the incredible contributions that Bob Williams and Michael Quinn Patton have made to the field of evaluation. Their books and texts provided me with endless guidance and insight, and I so admire their approach and outlook as evaluation *practitioners*.

I would like to thank those who provided financial assistance: the National Research Foundation, the UCT School of Management Studies, and the GeJuSTA Project.

To Teresa, thank you for your editorial support, watchful eye, and kind words of encouragement.

Last, I have to acknowledge myself. I would not have finished without the love and support of those above, but, ultimately, I did it. PhDs stretch on for such long periods that the researcher is bound to experience many lifetimes within it. I certainly did. Everything in my life has changed since I began this journey, but this thesis has remained constant. It took an extraordinary amount of determination to hold onto this thing while everything else seemed to collapse and rebuild around it. But I did it. And I am so proud of myself for doing so.

Table of Contents

Chapter One: Introduction	1
A Note on Terminology	2
Non-profit Organisation (NPO)	2
Monitoring, Evaluation, and Learning (MEL)	3
Systems Thinking	3
Background and Rationale for the Research	4
Research Objectives and Questions	7
Chapter Two: Monitoring, Evaluation, and Learning	9
The Basics: Programme MEL	9
Theory-based MEL	10
Management Orientations	11
Linear Approaches to MEL	12
MEL in South African NPOs: A Brief Historical Overview	15
The Problem with RBM	17
The Problem with Logframes and Pipeline Models	20
Chapter Conclusion	23
Chapter Three: Systems Thinking as a Theoretical Lens for MEL	24
What is Systems Thinking?	25
Systems Thinking Ontologies	28
The First Wave: Hard Systems	29
The Second Wave: Soft Systems	29
The Third Wave: Critical Systems	30
A Fourth Wave?	31
Systems Thinking and Evaluation	33
Uses of Systems Thinking in Evaluation	33

Barriers to the Use of Systems Thinking in Evaluation.....	38
The ‘Unapproachability’ of Systems Thinking.....	39
Wicked Problems.....	40
Overemphasis on Summative Evaluation	44
Chapter Conclusion	48
Chapter Four: ECD as an Interesting Test Case	50
The Importance of ECD	51
The Complexity of ECD	52
The Socio-ecological Model.....	52
The Knowns of ECD.....	58
Chapter Conclusion	61
Problem Statement and Research Questions	63
Chapter Five: Methodology.....	65
Toolkit Rationale	65
Research Phases	65
Research Paradigm	66
Research Design	69
Qualitative Orientation.....	69
Phase One: Toolkit Design and Expert Review.....	70
African Evaluation Principles and Made in Africa Evaluation.....	71
Phase One Sampling: Expert Reviewers	73
Procedure and Data Collection: Expert Review	75
Incorporating Expert Feedback	76
Phase Two: Testing the Toolkit.....	78
Phase Two Sampling	79
Test Cases	81

Phase Two Fieldwork and Data Collection.....	85
Phase Two Data Analysis.....	89
Ethical Considerations	96
Chapter Six: The Development of the SAMEL Toolkit	98
Description of the SAMEL Toolkit	98
Toolkit Audiences	99
Toolkit Design	100
SAMEL Conceptual Background (1): The Systems Concepts.....	102
Systems Thinking and MAE.....	102
Addressing Critique of Systems Concepts	105
SAMEL Conceptual Background (2): Theory-based Evaluation.....	107
Chen’s Action Model/Change Model Schema	109
Chen’s Schema and the SAMEL Toolkit.....	112
SAMEL Pragmatist Framework: Bricolage.....	112
Making Theoretical Decisions: A Reflexive Pause	114
Chapter Conclusion	116
Chapter Seven: Content of the SAMEL Toolkit.....	119
Expert Review	121
Toolkit Chapter 1: Introduction	121
Toolkit Chapter 2: Scoping.....	122
Evaluator Preparation for Scoping Workshop: Situation Analysis.....	125
Scoping Summary	136
Expert Feedback on Scoping Chapter	137
Toolkit Chapter 3: Focusing	138
Theory of Change	138
Theory of Action	154

Expert Feedback on Focusing Chapter	165
Toolkit Chapter 4: Monitoring & Learning	166
Monitoring.....	168
Learning.....	174
Expert Feedback on Monitoring & Learning Chapter	179
Toolkit Chapter 5: Evaluation.....	180
Evaluation Design	182
Evaluation Readiness	189
Expert Feedback on Evaluation Chapter.....	193
Expert Review of the SAMEL Toolkit: Overall Comments	195
Chapter Conclusion	198
Chapter Eight: Testing of the SAMEL Toolkit.....	200
Test Planning and Team Commitment.....	201
Scoping Workshop.....	203
Tool 1: Evidence.....	205
Tools 2 and 3: The SEM and Actor Map.....	205
Tool 4: Causal Loop Diagrams and System Archetypes.....	208
Tool 5: CATWOE and Root Definitions	211
Programme Scope	213
Focusing Workshops.....	215
ToC Workshop	215
Theory of Action Workshop.....	222
Monitoring & Learning Workshop	227
The Monitoring Plan.....	228
The Learning Strategy.....	231
Evaluation Workshop.....	232

The Evaluation Plan	233
Team Readiness	236
Workshop Aftermath	238
Heather’s Reflections	238
EduSpark.....	242
EYEN	244
Key Feasibility Issues.....	246
Facilitator application of concepts.....	246
Atmosphere.....	247
Rapport and Facilitation Style.....	247
Team dynamics.....	247
Time constraints	248
Chapter Conclusion	248
Chapter Nine: SAMEL Test: Thematic Analysis Findings	250
Thematic Overview.....	250
Theme 1: Strategic clarity (Programme Theory).....	250
Theme 2: Contextual disconnect.....	250
Theme 3: Organisational disconnect	251
Relationships between Themes.....	251
Discussion of Findings	253
Theme 1: Strategic clarity (Programme Theory).....	253
Theme 2: Contextual disconnect.....	258
Theme 3: Organisational Disconnect.....	264
Relationships between Themes.....	270
Chapter Ten: Discussion	273
The Approachability of the Toolkit for Evaluators.....	273

Systems Thinking’s Major Contribution to MEL: Programme Theory.....	277
The Shortcomings of Systems-based Programme Theories	281
The Incongruence of Systems Thinking and Accountability-led RBM.....	287
Intervention Design: Actioning the ToC.....	288
Becoming a Learning and Adaptive Organisation.....	293
How Feasible is a Systems Thinking Approach to MEL Development in NPO Contexts? Conclusions and Directions for Future Research	296
Limitations.....	301
Contribution of New Knowledge	302
References.....	305
Appendix A: Letters of Invitation to Expert Reviewers	332
International Expert.....	332
Local Experts.....	333
Appendix B: Reviewer Consent Form	334
Appendix C: Introductory Letter to Toolkit (Reviewers).....	335
Appendix D: Expert Review Form	336
Appendix E: Pilot Advert.....	340
Appendix F: NPO Eligibility Form	342
Appendix G: Letter to Participate in Pilot	345
Appendix H: Focus Group Schedules.....	346
Appendix I: Observation Schedules	351
Appendix J: Final Debriefing Interview Schedule- Heather	353
Appendix K: Follow-up Interview Schedule – NPO Directors.....	354
Appendix L: Final Codes.....	355
Appendix M: Theme Descriptions	361
Appendix N: Ethical Approval Letter.....	365

Approval of Extension.....	366
Appendix O: Data Management Plan	367
Appendix P: Workshop Participant Consent Forms	370
NPO Programme Team	370
Programme Stakeholders	371
Heather	372
Appendix Q: Life Space Example	373
Appendix R: Certificate of Editing.....	375

List of Figures

Figure 1. W. K. Kellogg Foundation Basic Logic Model showing If-Then Logic.....	13
Figure 2. Bronfenbrenner's Socio-ecological Model.....	53
Figure 3. Excerpt from Consolidated Expert Feedback.....	77
Figure 4. Excerpt from Actionable Feedback Matrix.....	78
Figure 5. EduSpark Profile.....	83
Figure 6. EYEN Profile	84
Figure 7. Illustration of Active Reading of the Data	92
Figure 8. Initial Mapping of Codes into Candidate Themes.....	93
Figure 9. Brainstorming Candidate Themes	94
Figure 10. Alignment Between MAE and Systems Concepts	104
Figure 11. The Action Model/Change Model Schema (Comprehensive Form)	110
Figure 12. SAMEL Toolkit Chapter Structure.....	120
Figure 13. Excerpt from Chapter 1 Explaining Systems Thinking Rationale	122
Figure 14. SAMEL Constituents of Programme Scope	123
Figure 15. Scoping Workshop: Summary of Sessions	124
Figure 16. SAMEL Situation Analysis Questions with Associated Systems Concepts and Tools	126
Figure 17. Example of a Vicious Cycle.....	131
Figure 18. Limits to Growth Example.....	133
Figure 19. Example of Root Definitions using CATWOE	136
Figure 20. ToC Elements and Associated System Thinking Tools	140
Figure 21. Function of Scoping Elements in ToC development.....	141
Figure 22. ToC Workshop: Summary of Sessions.....	142
Figure 23. Realist Context Surrounding the Programme	145
Figure 24. Realist Mechanisms	146
Figure 25. Life Space Model with Colour-coded Forces.....	149
Figure 26. Example of Matching Context to Mechanisms	151
Figure 27. Example of Reinforcing Feedback.....	152
Figure 28. Example of Unintended Feedback.....	153
Figure 29. Example of ToC Illustration using Non-linear Relationships.....	154

Figure 30. Example of a ToA Illustration	155
Figure 31. ToA Elements and Associated System Thinking Tools	156
Figure 32. ToA Workshop: Summary of Sessions.....	157
Figure 33. Reinforcing vs Balancing Effect of Programme Quality and Reputation on Participant Enrolment.....	162
Figure 34. Service Utilisation Diagram of the MomConnect Programme.....	164
Figure 35. Monitoring & Learning Workshop: Summary of Sessions	168
Figure 36. Toolkit Excerpt: SAMEL Monitoring Plan Summary	170
Figure 37. Toolkit Excerpt: Learning Questions using Systems Concepts	171
Figure 38. Toolkit Excerpt: SAMEL Learning Strategy Summary	175
Figure 39. The 3R Model	176
Figure 40. Meadows's Leverage Points Illustrated with SAMEL Reflective Questions	178
Figure 41. SAMEL Evaluation Workshop: Summary of Sessions	181
Figure 42. Summary of SAMEL Evaluation Criteria	185
Figure 43. Toolkit Excerpt: Example of Evaluation Questions for Relevance Criterion	186
Figure 44. Toolkit Excerpt: Adapted CSH Questions for Evaluation Bounding	188
Figure 45. The SAMEL Evaluation Boardgame's Gameboard.....	192
Figure 46. Heather Facilitating EYEN's Actor Map	207
Figure 47. Amended CLD Session: Listing and Linking Causes.....	210
Figure 48. Me Facilitating the CATWOE Discussion	212
Figure 49. Heather Neatens the EYEN ToC During Lunch. Middle Picture = Before; Right Picture = After	220
Figure 50. Example of a Life Space Shown to the EduSpark Team.....	221
Figure 51. EYEN Team Member Draws a Life Space Model	226
Figure 52. EYEN Team Members (Left) and EduSpark Team Members (Right) Play the Board Game	237
Figure 53. Four Non-linear Relationships between the Three Themes	252
Figure 54. Iceberg Model Summarising the Key Findings of this Research.....	300

List of Tables

Table 1. Logframe Structure	15
Table 2. Research Phases Aligned to Design Science Research Steps.....	66
Table 3. Pragmatic Beliefs that Underpinned this Research	68
Table 4. Summary of the African Evaluation Principles 2021	72
Table 5. Profile of Expert Reviewers	75
Table 6. Sensitising Concepts for Workshop Observation	87
Table 7. SAMEL Outline	99
Table 8. Applying Instructional Design Principles to the Toolkit	101
Table 9. Major Advantages of Scoping and Focusing Systems Thinking Tools	278

Acronyms and Abbreviations

AfrEA	African Evaluation Association
AEPs	African Evaluation Principles
CAS	Complex adaptive systems
CAQDAS	Computer-assisted qualitative data analysis software
CATWOE	Customers, actors, transformation, worldviews, owners, environment
CBO	Community-based organisation
CLD	Causal loop diagram
CMO	Context–mechanism–outcome
CREAM	Clear, relevant, economic, adequate, monitorable
CRISP	Clear, relevant, inclusive, efficient, pragmatic
CSH	Critical Systems Heuristics
CSP	Critical Systems Practice
DAC	Development Assistance Community
DBE	Department of Basic Education
DE	Developmental evaluation
DPME	Department of Planning, Monitoring and Evaluation
DSRP	Distinctions, systems, relationships, and perspectives
ECB	Evaluation capacity building
ECD	Early childhood development
EES	European Evaluation Society
ELOM	Early Learning Outcomes Measure
ELP	Early learning programme

FBO	Faith-based organisation
HSD	Human Systems Dynamics
M&E	Monitoring and evaluation
M&L	Monitoring and learning
MAE	Made in Africa evaluation
MEL	Monitoring, evaluation, and learning
NGO	Non-governmental organisation
NIECDP	National Integrated Early Childhood Development Policy
NPO	Non-profit organisation
OECD	Organisation for Economic Co-operation and Development
OM	Outcome Mapping
OMLS	Outcome Mapping Learning Community
PBO	Public benefit organisation
RBM	Results-based management
RMA	Reflexive Monitoring in Action
SAMEA	South African Monitoring and Evaluation Association
SEM	Socio-ecological Model
SETIG	Systems and Evaluation Topical Interest Group
SMART	Specific, measurable, achievable, relevant, timebound
SSM	Soft Systems Methodology
STCS	Systems thinking and complexity sciences
STS	Sociotechnical systems thinking
SWOT	Strengths, weaknesses, opportunities, threats
TBE	Theory-based evaluation

ToA	Theory of action
ToC	Theory of change
UNDG	United Nations Sustainable Development Group

Abstract

Recent trends in programme evaluation highlight increasing interest in using systems thinking to enhance the relevance and suitability of evaluation in complex social settings. However, the literature reveals a significant gap in practice-oriented empirical research on applying systems thinking to monitoring, evaluation, and learning (MEL) in non-profit organisations (NPOs). This research addresses this gap by testing the feasibility of a systems approach to MEL development in NPO settings. Employing a multi-phase exploratory design, I developed the Systems Approach to MEL (SAMEL) Toolkit for evaluators to aid NPOs in developing MEL frameworks through structured, facilitated workshops. The Toolkit was first reviewed by experts and adjusted accordingly. It was then tested with two early childhood development NPOs, with workshops facilitated by myself and another evaluator. Data were collected using workshop observations, focus group discussions, interviews with the evaluator, facilitation journals, and follow-up interviews with NPO directors and the other evaluator. The data were analysed using reflexive thematic analysis. Key findings suggest systems thinking can significantly enhance strategic planning and MEL readiness in NPOs through its application to programme theory. However, MEL implementation faces substantial challenges due to contextual and organisational complexities, particularly with regard to funding. While systems thinking enhances NPO teams' strategic clarity and critical appraisal of programme design, NPOs have limited ability and capacity to act on these insights. An accountability culture and results-based management orientation to funding significantly constrains the flexibility demanded by systems thinking approaches. Based on these findings, this thesis recommends exploring ways to enhance the suitability of systems thinking for NPOs, including capacity-building and changes to rigid funding and reporting requirements.

Keywords: monitoring, evaluation, and learning; non-profit organisations; systems thinking; South Africa

Chapter One: Introduction

In both developing and developed nations across the world, social interventions and programmes play a critical role in the response to societal needs and filling gaps in state provision for vulnerable populations. Sometimes, these responses are rapid and crisis-driven, while at other times, they are long-term solutions to historical and deeply entrenched inequalities. What can be argued to be universal, however, is that both the act of social intervention and its results are messy, complicated, dynamic, and, oftentimes, unpredictable, which is because human beings and globalised societies are messy, complicated, dynamic, and unpredictable.

This complexity has perhaps never been so plainly observable and widely talked about as in recent times as the world, collectively, faces and tries to recover from increasing global hazards such as climate disasters and the COVID-19 pandemic. Michael Quinn Patton (2021), an avid advocate for transforming evaluation to meet these growing and urgent global changes, argues:

The pandemic has been global in scale and universal in impact. The climate emergency will, likewise, be global in scale and universal in impact. The global climate emergency affects all of us, leading to calls for action in whatever niche we inhabit. For evaluators, that niche is evaluation. Evaluation has emerged as critically important in realizing the vision and aspirations of the sustainable development goals. But current evaluation criteria and practices are inadequate for evaluating transformation. That makes evaluation part of the problem. Evaluating transformation means transforming evaluation. (p. 55)

We arguably face an ongoing ‘polycrisis’ –multiple, overlapping and interconnected crises that deepen social inequities and threaten the health of our planet (Buckton et al.,

2024; Patton & Richardson, 2024). As such, many have echoed Patton's calls for transformation, urging evaluators to embrace systems thinking in their evaluation practice to engage with the complexity of the polycrisis (Buckton et al., 2024). This involves understanding evaluands and their contexts from a systemic perspective, which, in turn, enables evaluators to address the root causes of problems more meaningfully, engage multiple perspectives on social issues and interventions, incorporate social justice and indigenous knowledge systems, and support continuous learning and adaptive practices in dynamic contexts (Buckton et al., 2024; Gates et al., 2021; Patton & Richardson, 2024; Uitto, 2021; Walton & Baker, 2021).

This thesis represents my personal response to this call, undergirded by my own professional experiences in monitoring, evaluation, and learning (MEL) consulting with non-profit organisations (NPOs), as well as my personal interest in systems thinking. Before I describe the role of this experience in shaping this thesis, I will briefly delineate key terms used throughout this thesis.

A Note on Terminology

Throughout this thesis, I define and use several terms, but none as important or ubiquitous as (1) *NPO*, (2) *MEL*, and (3) *systems thinking*, necessitating their acknowledgement upfront.

Non-profit Organisation (NPO)

Organisations in South Africa's civil sector go by several different acronyms, including NPO (non-profit organisation), NGO (non-governmental organisation), CBO (community-based organisation), FBO (faith-based organisation), and PBO (public benefit organisation). Prior to 1994, many organisations were established to oppose the apartheid government and provide services to oppressed communities, and were thus staunchly non-governmental in

structure and status (Abrahams, 2015). Post-1994, as the country transitioned to a democratic state, civil society organisations began to work with, or were funded by the government, and had to apply for non-profit status (Abrahams, 2015). South Africa's Nonprofit Organisations Act 71 of 1997, provides legal standing to NPOs, together with a voluntary registration mechanism. As such, the Act recognises NPOs as legal entities, and to register as an NPO, the organisation can be an NGO, FBO, or CBO. I use the term *NPO* throughout the thesis to refer to civil society organisations operating in South Africa's third sector, irrespective of whether they are formally registered.

Monitoring, Evaluation, and Learning (MEL)

MEL is an interdisciplinary field that goes by many different names. The term *monitoring and evaluation* (M&E) is most pervasive, referring to an organisational management system established to (a) monitor a programme's implementation and progress towards its goals and objectives over time, and (b) to systematically evaluate the programme, to inform decision-making (Görgens-Albino & Kusek, 2009). The addition of the *L* (learning) has become more prominent in recent years as practitioners and participants recognise the critical importance of learning from M&E data and evidence. I subscribe to this notion, and designed a large portion of this research to reflect a learning orientation. For this reason, I use the term *MEL* throughout this thesis.

Systems Thinking

This thesis grapples with this term in great depth and breadth, but I approach this large and nebulous discipline in a broad and collective sense by referring to it as *systems thinking*. Systems thinking is generally considered a discipline that is concerned with a whole system, which can be defined as a set of interconnected elements that are organised in such a way as to achieve a particular purpose (Arnold & Wade, 2015; Meadows, 2009). Some

systems thinkers are located within particular scientific disciplines (e.g., engineering), and thus focus on particular systems theories aligned to that discipline (e.g., complex adaptive systems, complexity theory, general systems theory, or system dynamics). Others take a more generalist approach, applying systems thinking more broadly and in a less discipline-specific manner. For example, Cabrera et al. (2008) define systems thinking as “a pattern of thought”(p. 306) or the application of ‘systems’ to any discipline or body of knowledge that informs and frames our understanding of the world (Rajagopalan & Midgley, 2015). I am of the view that systems thinking is valuable in a myriad of domains, and I use it as a lens through which to view and engage with MEL. Particular theories and approaches are presented along the way, but the term is used throughout to refer to a corpus of methods, tools, concepts, theories, and ideas that could provide a way in which to view the world.

Background and Rationale for the Research

This thesis was shaped by my own academic and professional experience as an MEL consultant. For eight years now, I have worked primarily with NPO teams to develop their MEL strategies and conduct evaluations, and for six of those years, my clients have operated mainly in the early childhood development (ECD) sector.

ECD is a fascinating and challenging space to work in, and my consulting experience has been characterised by both awe and frustration — awe at the dedication, creativity, and tenacity shown by NPOs in advocating for children, and frustration with the MEL and evaluation tools available to me to support these teams adequately and in the most sensible way possible. There has always been a feeling that the MEL and evaluation products that I develop are missing the mark in some way.

With an undergraduate degree in psychology and organisational psychology, I became particularly interested in the organisational development of NPOs and the role played by the

people doing the work in achieving the results. In January 2020, I considered pursuing a doctoral thesis on the role of organisational effectiveness in programme effectiveness. I came across two books that were very influential: Peter Senge's *The fifth discipline: The art and practice of the learning organization* and Jamshid Gharajedaghi's *Systems thinking: Managing chaos and complexity*. These works introduced me to a completely new way of thinking — in systems.

Just three months later, the world changed instantly, swept into rapids of change due to the COVID-19 pandemic. Life was suddenly strange and unpredictable. The term *unprecedented* became *precedented*, and few things were as certain as the uncertainty of the moment. Wealthy communities stockpiled toilet paper and packed their freezers, while the most vulnerable populations were left bereft. My ECD clients were trying to keep poor children protected and learning in their homes as they faced the effects of the lockdown on the already entrenched inequities in young children's development (Ebrahim et al., 2021; Spaul & Van Der Berg, 2020).

As captured in Patton's (2021) quote earlier, the crisis of the pandemic highlighted just how fragile, unpredictable, and complex social intervention is. Swift, adaptive action in an ever-changing context became imperative for organisational and community survival. As ECD NPOs shifted their core function from supporting early learning teachers to feeding starving communities, I wondered — with a hefty dose of cynicism — how well the linear, input–output logic models that undergird so many MEL frameworks were faring.

While pandemics might represent the extreme end of a spectrum of life and society's complexity, it became clear to me what was missing in my work with NPOs: big-picture thinking — not just in terms of scale but also in depth, connection, and perspective.

As I learned about systems thinking, I saw how fitting the discipline is for understanding both social intervention and evaluation. This was not a new revelation; a systemic view had been in the evaluation lexicon for decades at that point (e.g., Williams & Imam, 2006). Authors had, for some time, been urging civil society, including NPOs and funders, to relinquish simple, mechanistic, and linear approaches to intervention strategy and to replace them with a more adaptive learning strategy – one that can better respond to the uncertainty and unpredictability that characterise more complex social problems (e.g., Kania et al., 2014; Patrizi et al., 2013).

Traditional evaluation and programme monitoring approaches tend to focus on linear approaches to understanding and assessing social interventions. Theories of change, logic models, and logical frameworks typically depict programme pathways as a linear causal sequence, beginning with the programme inputs and activities and ending in participant outcomes and impacts (Reynolds et al., 2016).

However, social problems are rarely stable and linear, particularly in the current landscape; they are adaptive and interconnected, and subject to multiple differing perspectives on what they are and how to solve them (Gates, 2016). As such, linear approaches to understanding and evaluating social programmes are argued to be too rigid and static to be able to support the adaptive learning needed in the realm of social programming (Bornstein, 2003, 2006; see also Kushner et al., 2021), particularly as they fail to adequately account for a programme's context and interactions with other variables, which limits the interpretation of programme effects (e.g., Davies, 2004; Dyehouse et al., 2009; Renger et al., 2019).

While the problems are well documented, with a myriad calls for transformation, how to effect the transformation is less clear. Many evaluators have historically, and still currently,

advocate for incorporating both complexity science and systems thinking into evaluation, in a bid to transform the discipline (e.g., Gates et al., 2021; Hummelbrunner, 2011; Patton, 2011; Williams & Imam, 2006). A major identified barrier, though, is evaluators' lack of familiarity with these disciplines and how to apply them in their work (Walton, 2016). This has been addressed over time, with the publishing of guidance documents and evaluation examples that centre on the design of summative evaluations guiding external evaluators who are conducting outcome or impact evaluations of complex programmes, or those aimed at changing systems (Hargreaves, 2010; Kehrer et al., 2020; Latham, 2014; Parsons, 2007; Preskill et al., 2014; Stephens et al., 2018; Trochim et al., 2016; Wasserman, 2010; Williams & Imam, 2006). However, far less guidance is available to guide the development of MEL frameworks, which, I argue, is a particularly important component in the support of adaptive social programming in a complex world.

For reasons unknown, MEL as a function of social programme management has largely been neglected in the literature on systems thinking and evaluation. There is a lack of widely available guidelines or resources to support MEL practitioners working with NPOs to develop systems thinking-informed frameworks. This thesis thus aims to meet this gap.

Research Objectives and Questions

In order to fill the gap in available guidelines on systems thinking and MEL, this research was designed to meet the following objectives:

1. Explore systems thinking methodologies and design a systems thinking-based toolkit that can guide MEL framework development in a participatory workshop format;
2. Test the designed toolkit amongst South African ECD NPOs.

The associated research question was: How feasible is a systems thinking approach to MEL development?

The sub-questions were:

1. How did the workshop participants and the evaluators (workshop facilitators) experience the Systems Approach to MEL Toolkit as a systems thinking approach to MEL?
2. What are the challenges in and advantages of integrating systems thinking into MEL?

Chapter Two: Monitoring, Evaluation, and Learning

Sitting in on NGO meetings ... I was often struck by the disdain with which staff greeted discussions about targets, indicators and impact measurement. 'If we can't give the numbers we don't get the money' was a complaint often voiced [...] The perception of M&E as a 'tick-box exercise' and as a mechanism for uniformity conveyed a self-portrayal of NGOs as weak, with no agency and little power in hierarchical reporting structures. (Mueller-Hirth, 2012, p. 10)

Mueller-Hirth's reflection on the observed relationship between NPOs and MEL illustrates a pervasive frustration with rigid, donor-driven performance metrics that reduce complex social impacts to simplistic numerical outputs. I have seen this sentiment often in my work as an MEL consultant and have expressed it myself.

This chapter critically examines the current practices and challenges of MEL in NPOs. It begins by outlining the foundational principles of MEL and its adoption in South Africa, and then provides a critique of prevalent accountability-based and linear approaches to MEL, particularly results-based management (RBM) and logical frameworks. Through a discussion of literature, this analysis aims to demonstrate the necessity of evolving MEL practices in order to enhance its role in organisational learning. The chapter concludes with a particularly promising pathway for change — systems thinking, which is explored further in Chapter Three.

The Basics: Programme MEL

Programme evaluation is the use of social science research methods to assess social programmes through the investigation of the programme's conceptualisation of the problem, the way that it was designed and implemented, the outcomes it achieved, and its efficiency (Rossi et al., 2019). Programme monitoring involves the systematic collection of programme

performance data that indicate whether the programme is being implemented as it was intended and whether it is achieving the intended outcomes (Rossi et al., 2019).

Monitoring is primarily an internal activity that serves programme management and accountability, whereas evaluation findings inform decision-making about the programme, its strategic direction, and future investment (Peersman et al., 2016). Monitoring data, however, can also be used to support evaluations (Markiewicz & Patrick, 2016). Both monitoring and evaluation activities are typically guided by an MEL framework, which is a planning tool that outlines the routine monitoring and periodic evaluation activities that are to be conducted during a programme's lifespan (Markiewicz & Patrick, 2016). MEL frameworks are an important mechanism for establishing and tracking programme results, as well as gathering data to support organisational learning, strategic decision-making, and programme improvement (Markiewicz & Patrick, 2016).

Common constituents of MEL frameworks include: a problem or situation analysis; steps for user engagement; programme theory; evaluation questions; indicators; targets; data collection instruments and data sources; responsibilities and time frames; data collection, management, and analysis guidelines; and reporting, learning, and dissemination strategies (Kusek & Rist, 2004; Markiewicz & Patrick, 2016).

Theory-based MEL

Programme theory proposes a model of how a programme is supposed to work, i.e., how each of its activities or interventions are expected to result in outcomes and make an observable change amongst the programme's target population (Funnel & Rogers, 2011). Theory-driven, theory-based, or theory-guided evaluation involves the development of a programme theory that models how the particular programme is expected to bring about

change. This model is then used as the conceptual framework for assessing the programme's effectiveness (Chen, 2015).

Basing evaluation on programme theory opens up the 'black box' of method-driven evaluation to uncover *why* a programme works (not just whether it works) through careful attention to cause-and-effect relationships (Chen & Rossi, 1983; Weiss, 1997). Programme theory can incorporate established social science theory as the basis for programme theory, or it can articulate programme stakeholders' assumptions and implicit models for how programmes are expected to work (Chen, 2015; Rossi et al., 2019). Once the theory has been established, evaluators can develop and prioritise evaluation questions and design the evaluation around answering them (Donaldson, 2021).

Programme theory serves the same function in MEL by guiding the identification of priority areas for measurement and monitoring (Funnel & Rogers, 2011; Görgens-Albino & Kusek, 2009; Markiewicz & Patrick, 2016). Conceptual tools often used to support theory-based MEL are the theory of change (discussed in Chapter Five), the logic model, and the logical framework (the latter two are discussed below).

In sum, theory-based MEL provides a clear framework for MEL design and implementation, uncovering the assumed causal relationships that illuminate how and why programmes are expected to work and, thus, what needs to be measured on an ongoing basis.

Management Orientations

There are several management orientations that can guide the application and utility of MEL frameworks. RBM is a management strategy that was initiated in the 1990s by the United Nations to improve effectiveness and accountability (United Nations Development Group [UNDG], 2011). RBM is aimed at (1) aligning all programme processes, products, and

services with the achievement of intended results, (2) monitoring actual results, and (3) using this monitoring data to inform decision-making, planning, and accountability (UNDG, 2011). A key strength of the RBM approach is linking MEL with planning (Markiewicz & Patrick, 2016).

Adaptive management, which has its roots in conservation and natural resource management, is a strategy of using evidence to adapt programmes in conditions of ongoing uncertainty, where traditional planning is not possible (Rogers & Macfarlan, 2020). In uncertain situations, collecting implementation- and outcome data makes the unseen seen. MEL functions and data are thus central to this approach and provide the basis for programme implementation and management (Rogers & Macfarlan, 2020).

Linear Approaches to MEL

As noted earlier, the logic model and logical framework are two MEL tools commonly used to explicate the assumptions underlying a programme and design measurement systems to track programme performance.

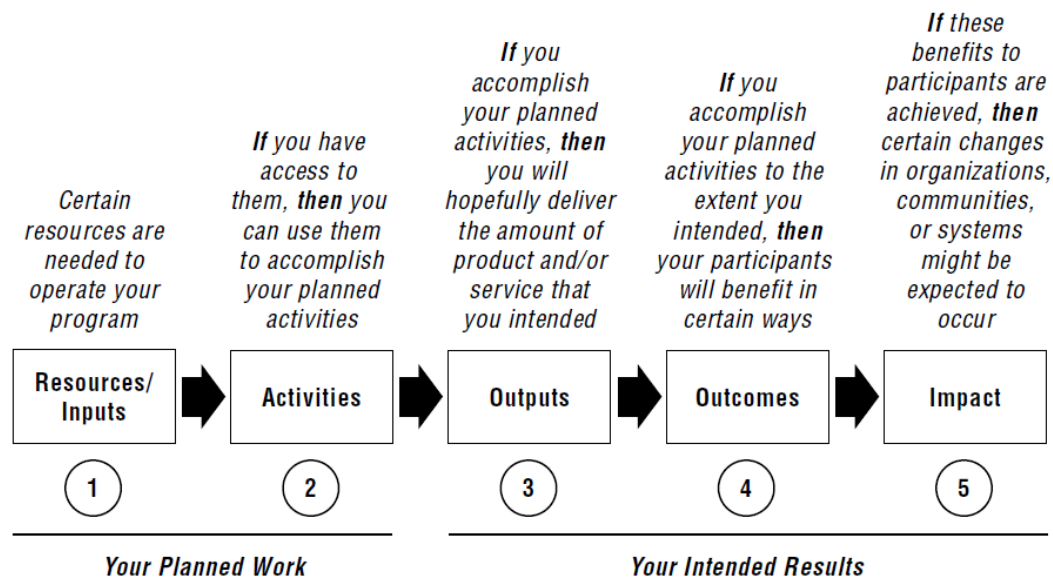
The logic model horizontally depicts a programme's inputs or resources, activities, outputs, outcomes, and impacts, as shown in Figure 1 (Donaldson, 2007; Funnel & Rogers, 2011; Knowlton & Phillips, 2013; Wholey, 1979).

Here, the term *logic* refers to the theory component — the causal assumptions underlying a programme's design. Funnel and Rogers (2011) refer to these models as 'pipeline logic models', as they illustrate programme processes as linear, with inputs at one end and outputs and outcomes at the other. As can be seen in

Figure 1, logic models employ an if–then logical sequence between the components in each column; if the information in one column is true, the result in the next column will be achieved (Knowlton & Phillips, 2013).

Figure 1

W. K. Kellogg Foundation Basic Logic Model showing If-Then Logic



Note. From “Logic Model Development Guide”, by W. K. Kellogg Foundation, 2004, p. 3

(<https://wkkf.issuelab.org/resource/logic-model-development-guide.html>). Copyright 2004 by W.K. Kellogg Foundation.

In the 1990s and early 2000s, logic models were taken up by international donor agencies such as United Way of America and the W.K. Kellogg Foundation, who published development guides in the late 1990s and early 2000s (see United Way of America, 1996; W. K. Kellogg Foundation, 2004). This shift reflected growing pressure in the United States (within governmental and non-profit sectors) to demonstrate programme results using results-oriented management (Wholey, 2001). Results-oriented management involves

developing programme strategies and outcomes, measuring programme performance (with an emphasis on outcomes), and then using the performance data to inform programme improvement (Wholey, 2001). The logic model served as a particularly useful tool in this context; it could be used to surface stakeholders' assumptions about how the programme works and create a shared understanding of programme theory. The logic model guided programme design and planning by helping stakeholders think about the key programme components and goals, and it provided a framework for the monitoring of both programme implementation and outcomes (W.K. Kellogg Foundation, 2004).

While some refer to logical frameworks as a type of logic model (e.g., Funnel & Rogers, 2011), logical frameworks (or 'logframes') were developed separately from logic models and theory-based evaluation. Logframes were first developed and adopted by global donors in the 1970s as a programme planning and tracking tool (Wildschut, 2014). While the developers were also concerned with the cause-and-effect relationships within a programme (as with logic models), the overall concern was with programme management, accountability, evaluation, and reporting (Wildschut, 2014). The major output of the logframe development process is a tabular matrix that describes a programme's underlying logic in an if-then manner; i.e., if Activity A is implemented, then Outcome X can be expected to result (Ringhofer & Kohlweg, 2019).

Table 1 displays the typical matrix format: for each model component (impact, outcomes, outputs, and activities), the matrix describes the component and provides objectively verifiable indicators and means of verification, as well as assumptions. While indicators and means of verification provide a measurement framework to track each of the programme components, the *Assumptions* column contains contextual information on the factors (inside and outside of the programme) thought to affect the realisation of programme results (Funnel & Rogers, 2011).

Table 1

Typical Logframe Structure/Matrix

	Narrative description	Indicator	Means of verification	Assumptions
Impact				
Outcomes				
Outputs				
Activities				

Given their standardised structure, logframes offer donors an easy way to track programme results. Also used in pursuit of results-oriented or results-based management, logframes are tailored by individual organisations to their own philosophies and approaches to learning and management (Wildschut, 2014). They offer a simple overview of a programme that can aid planning and management by programme teams. Today, logframes form the dominant approach to programme design, monitoring, and evaluation globally, including in South Africa (Ringhofer & Kohlweg, 2019; Wildschut, 2014).

MEL in South African NPOs: A Brief Historical Overview

MEL rose to prominence in South Africa in the 1980s and 1990s, as NPOs were required to evaluate their activities and outcomes to qualify for donor funding (Abrahams, 2015; Mouton, 2010). Mouton (2010) documents the rise of programme evaluation in South

Africa, noting that the international donor community established the discipline in the NPO sector. During apartheid, solidarity fundingⁱ meant that NPOs were allowed great flexibility, and monitoring and evaluation efforts were not rigorous and systematic. The most common tool used at the time was the logframe (Mouton, 2010).

Following South Africa's transition to a democracy in 1994 (and when NPOs themselves transitioned from apartheid activists to more formal service providers), donors began to place greater emphasis on accountability (Mouton, 2010). Influenced by new managerialism and New Public Managementⁱⁱ, international funders and the South African government paid greater attention to the need for performance measurement and accountability, subsequently shifting towards an RBM orientation (Abrahams, 2015; Mouton, 2010). In the early 2000s, the South African presidency established the Department of Performance (later 'Planning'), Monitoring and Evaluation (DPME) and the country's first related professional association, the South African Monitoring and Evaluation Association (SAMEA) (Abrahams, 2015). This led to enhanced evaluative capacity-building in the country, with universities offering formal evaluation training programmes (Abrahams, 2015; Beney et al., 2015).

Despite these advancements, the MEL sector in South Africa faces challenges such as a mismatch between the supply of skilled evaluators and the demand for their expertise

ⁱ Financial support provided by international organisations, foreign governments, and global civil society groups to South African NPOs that were engaged in the anti-apartheid struggle. This funding was not merely financial assistance; it was an expression of solidarity and political support for the efforts to dismantle apartheid and promote human rights, social justice, and democracy in South Africa.

ⁱⁱ New Public Management is a management philosophy used by governments since the 1970s to modernise and render public sector services more efficient and similar to those of the private sector. It emphasises the use of business-like practices and performance metrics such as cost-efficiency, accountability, and service quality (Gruening, 2001).

(Blaser Mapitsa & Khumalo, 2018), as well as an overreliance on RBM (Bornstein, 2006; Kachur et al., 2016; Mueller-Hirth, 2012). The latter is of particular relevance to this thesis and is explored further below.

The Problem with RBM

South African NPOs are highly reliant on funders to keep them financially afloat (Maboya & McKay, 2019; Vorwerk Marren et al., 2024). The nature of the funder–NPO relationship has significant implications for the role of MEL in NPOs, given the interplay between financial dependence and RBM.

Resource dependence theory posits that organisations are not self-sufficient; they rely on external resources to survive (Pfeffer & Salancik, 2009). This dependence on external resources influences their behaviour and strategies, including their structure, decision-making processes, and interactions with other organisations. For example, NPO collaboration is limited because they compete for grants and funding (Kumaran et al., 2012). According to resource dependence theory, the key to organisational survival is the ability to secure and maintain essential resources, which can often result in power imbalances between organisations and the entities that control those vital resources (Pfeffer & Salancik, 2009). Reith (2010) critically analysed ‘partnerships’ between funders and NPOs, and argued that NPOs’ financial reliance on funders created a significant power imbalance.

These power dynamics can influence the way NPOs shape their programmes, as well as the ways in which MEL is perceived and implemented. NPOs’ financial dependence on donors can shape their strategic priorities and programme design. Funders' mandates and related grants frequently dictate the goals and structure of interventions, requiring NPOs to align their programme designs and activities with donor-defined objectives rather than context-specific needs (Bornstein, 2003; Reith, 2010; Witesman & Heiss, 2017; Yeboah, 2022).

This dynamic fosters strategic conformity, where the imperative to secure and maintain funding often supersedes innovation or responsiveness to local needs.

Secondly, MEL is an important governance tool for NPOs to secure and maintain funding. MEL and programme evaluations have arguably served as upward accountability mechanisms, where NPOs can maintain funding relationships by demonstrating to funders a 'return on their investment' (Eyben, 2013; Kelly, 2021; see also Lennie & Tacchi, 2013; Lowe & Wilson, 2017; Mueller-Hirth, 2012). This has important implications: (1) MEL is predominantly seen as an accountability tool rather than a means of learning and effective programme management (Bornstein, 2003, 2006; Kachur et al., 2016); and relatedly, (2) NPOs may thus use MEL inappropriately in a bid for financial survival (Lowe & Wilson, 2017; Yeboah, 2022).

To understand the implications of accountability-driven MEL, it is important to first consider organisations' evaluation policy - the explicit or implicit rules and principles that govern organisational practices and decisions regarding evaluative activities, and/or MEL in this case (Trochim, 2009). Trochim (2009) emphasises that evaluation policies profoundly shape evaluative behaviour, influencing how evaluations are conducted, who participates, and how findings are used by the organisation. When evaluation policies are externally imposed and fail to consider the organisational context, they can lead to misalignment with the organisation's needs, reducing the utility and sustainability of evaluation efforts.

I argue that RBM and accountability culture play a significant role in hindering effective evaluation policy and the institutionalisation of MEL in organisations. This is because MEL is framed as an externally-driven compliance and accountability tool, enabling NPOs to view MEL as a transactional or symbolic process, downplaying its role in learning and strategic

development, and thus deprioritising the development of evaluation policy, including the internal skills or systems that support sustainable MEL (Carman, 2011).

For example, RBM has been criticised for encouraging organisations to adopt short-term, output-focused, and inflexible MEL practices instead of practices that measure long-term impact in communities (French et al., 2023; Simister, 2017; see also Bevan & Hood, 2006; Goncharenko, 2021; Ríos Romero et al., 2023; Urwin et al., 2023). MEL systems are thus designed to track these ‘easy wins’ to validate the continuation of funding and account to funders, instead of accounting to communities regarding the change the programmes are designed to make (Essel, 2021; Eyben, 2013; Urwin et al., 2023). Furthermore, a funder-centric approach to MEL has been argued to foster ‘gamification’, due to financial “incentives for deception” (Bornstein, 2006, p. 53). This can involve manipulating programme implementation in order to achieve the results that funders expect to see (Lowe & Wilson, 2017; Yeboah, 2022).

RBM approaches such as the logframe are therefore argued to stifle the organisational learning that MEL can offer NPOs (Bornstein, 2006; Kachur et al., 2016). The need to rigidly adhere to pre-approved matrices of outputs and outcomes within specified time frames is “yet another expression of donor control over how projects are run, and failure to acknowledge their real-life complexity” (Reith, 2010, p. 452). Qualitative research conducted in South Africa (Kachur et al., 2016; Mueller-Hirth, 2012) found that NPOs perceive MEL primarily as rigid and burdensome accountability practices imposed by donors (‘ticking boxes’), rather than as a means of internal learning and improvement. Similarly, a qualitative study with organisational leaders across Africa argues that the misalignment between donor priorities and organisational contexts not only undermines the relevance of MEL tools but

also discourages the development of internal skills and systems needed for sustained evaluation capacity (Masvaure & Fish, 2022).

Thus, by having to implement externally-imposed MEL tools that are deemed contextually inappropriate or burdensome, organisations can lose trust in and motivation for MEL, key elements for embedding evaluative practices into organisational culture (Carman, 2011; Jacob, 2023; Wade & Kallemeyn, 2020). Wade and Kallemeyn (2020) emphasise that leadership and buy-in are critical to sustaining evaluative practices within organisations. Leaders who value and see the benefits of MEL advocate for it, allocating resources, and embedding it into organisational structures. Given that organisational factors, such as time and financial limitations, are known to hinder MEL in organisation (Nielsen et al., 2024), the role of buy-in from leadership becomes crucial for the successful institutionalisation of MEL and the development of evaluation policy.

In summary, RBM's funder-centric approach leaves little room for organisations to cultivate MEL advocates or integrate MEL into their strategic priorities. The externally imposed nature of some MEL tools, often designed without consideration for the organisational context, further undermines their relevance and perceived value. As a result, organisations frequently view these tools as burdensome and disconnected from their operational realities, reducing staff buy-in and ownership over MEL. This dynamic reinforces a focus on producing evidence to satisfy funders, perpetuating a cycle of dependency and preventing the institutionalisation of MEL as an integral part of organisational culture and decision-making.

The Problem with Logframes and Pipeline Models

They are Rigid and Burdensome. In the early 2000s, Bornstein (2003, 2006) published two articles on South African NPOs' experiences of monitoring and evaluation (M&E) and

donor relations. Interviewees in both articles criticised donors' enforcement of logframes and the management style that results from their usage, calling logframes "technicist", "sterile", and "reductive". These rigid and structured reporting tools cannot adequately communicate the messy and dynamic happenings of community-based change, yet programmes need to align with milestones and fit neatly into the boxes of the logframe's matrix. This makes it impossible to capture complexity within the context of the programme, and impedes flexible, responsive, and iterative programming (Bornstein, 2003). This is because logframe matrices are centred on the programme and its achievements, as opposed to the greater change pathways amongst the target population and their communities (Ringhofer & Kohlweg, 2019).

Wildschut's (2014) doctoral thesis on programme theory, conducted on South African NPOs, established that nearly half of the study's sample (47%) used logframes. Aligned with Bornstein's (2003, 2006) findings, Wildschut's (2014) survey of NPOs found that the majority find these reporting structures were not worth the time and are unsuitable for development work. A lack of perceived utility means that these tools tend to be regarded as static, inflexible models instead of being updated and continuously reflected upon as the programme evolves (Kushner et al., 2021). Newer studies on the use of logframes in South Africa are unavailable. However, based on my consulting experience with local NPOs, they are still ubiquitously used and required.

They are Reductionist. Despite the prevalence of logic models in programme planning and management, many evaluation authors oppose their use in complex contexts.

Kushner et al. (2021) make several arguments against linear logic models. First, the authors argue that logic models depict all programme components (e.g., inputs, activities, outputs, and outcomes) as weighted equally, which may be helpful for planning but detracts and distracts from understanding *why* a programme works. Programme assumptions are also

argued to be kept at the periphery of the programme theory, instead of being integrated into the identified causal relationships. Reynolds et al. (2016) take this point even further by arguing that “rigid input-output models” (p. 663) are, by definition, unable to accommodate ongoing and adaptive change, thus reflecting the sentiments shared by South African NGOs on logframes in Bornstein’s (2003, 2006) studies. Furthermore, the “plug and chug approach to filling in boxes” (Kushner et al., 2021, p. 6) limits an in-depth engagement with the complex reality of dynamic social change.

In support of these views, other authors have argued that simple linear models fail to depict interactions between outcomes and other variables, which can limit our understanding of the factors that influence the achievement of outcomes and result in an incomplete interpretation of programme effects (Davies, 2004; Dyehouse et al., 2009; Renger et al., 2019; Rogers, 2008). Likewise, Cabrera et al. (2008) argue that programmes do not operate in a vacuum, and investigating the relationships between all programme components and the context is important in strengthening evaluations.

These authors highlight the inability of simple linear models to capture a rich and contextual programme theory that communicates the change pathway adequately, and while these criticisms are aimed at logic models, it stands to reason that they also apply to logframes. Many of these authors offer alternatives or additions to existing formats to depict programme theory, using systems thinking (e.g., Dyehouse et al., 2009; Hummelbrunner, 2011; Kushner et al., 2021).

Systems thinking has, for some time, been proposed as an alternative conceptual lens through which evaluators and programme stakeholders can gain a more thorough understanding of programme effectiveness and community impacts (e.g., Gates et al., 2021; Patton, 2011; Ulrich, 1988; Williams & Imam, 2006). In the following chapter, I describe this

growing movement and the ways in which it can overcome the shortcomings of linear MEL models.

Chapter Conclusion

This chapter examined the evolution and application of MEL approaches in NPO settings, with a particular emphasis on traditional RBM and linear models such as logframes and logic models, illustrating their widespread adoption and subsequent criticisms. These traditional methods, once favoured for their structured, output-focused understanding of progress, are increasingly seen as rigid and reductionist. Critics argue that they often serve more as tools for securing funding by demonstrating accountability, rather than fostering genuine programme improvement and learning.

In conclusion, while traditional MEL frameworks have provided a foundational approach, there is a growing recognition of their limitations in understanding and addressing complex social issues. The chapter lays a foundation for an argument that I will expand upon in the following chapter: MEL needs to evolve to prioritise strategies that prioritise learning and a deeper and more nuanced understanding of social change, thus better supporting NPOs in achieving sustainable and impactful outcomes in their work. For decades, evaluators have advocated systems thinking as the catalyst for the required change.

Chapter Three: Systems Thinking as a Theoretical Lens for MEL

Social programs are undeniably, unequivocally, unexceptionally social systems. They comprise, as with any social system, the interplays of individual and institution, of agency and structure, and of micro and macro social processes. (Pawson & Tilley, 1997, p. 64)

This chapter delves into systems thinking as a pivotal conceptual lens through which change can be understood and, thus, through which MEL could be designed. It traces the evolution of systems thinking from its inception as a response to scientific reductionism to its current application across various fields. Central to this discussion is the argument that systems thinking, incorporating principles from complexity theory, provides a comprehensive framework that captures the multifaceted and interconnected nature of social phenomena, distinguishing it from traditional linear approaches to research and evaluation, which may oversimplify cause-and-effect relationships. I hope to persuade the reader to agree with Pawson and Tilley's (1997) assertion that programmes are systems, as are the environments in which they are implemented. We, therefore, need to develop systems thinking competence amongst both implementing organisations and evaluators.

The chapter is structured around the three historical 'waves' of systems thinking: hard systems, soft systems, and critical systems, each reflecting shifts in focus and inquiry methodologies that are relevant in applying systems thinking in programme evaluation. This chapter also presents an overview of the use of systems thinking in programme evaluation, highlighting the immense utility and relevance of the intersection of the fields and critically examining the barriers to implementing systems thinking in MEL practices. By integrating theoretical insights with practical application, the chapter advocates embracing systems

thinking in MEL, but emphasises the need for innovative and approachable resources for evaluators still absent in this field.

What is Systems Thinking?

The term *systems thinking* was reportedly first coined in 1987 by Barry Richmond (Arnold & Wade, 2015), who later said he liked where the emphasis was being placed — “on the thinking!” (Richmond, 1994, p. 139). Richmond (1994) proposed a dual understanding of the term: that it represents both a paradigm (a way of seeing the world) and a learning method (processes, language, and technologies used to study systems).

As a paradigm, systems thinking aspires to take a holistic view of social phenomena. The field first emerged in the 1940s, in reaction to the dominant paradigm of the time, scientific reductionism, whereby phenomena were understood by “decomposing” (Agazzi, 1978, p. 350) them into their smallest possible parts. Researchers framed inquiries by “looking down from the whole to the parts” (Agazzi, 1978, p. 350), meaning that the behaviour of the whole phenomenon is understood through the behaviour of the parts. Instead of concentrating on small, single variables, systems thinking encourages consideration of the connections, relationships, patterns, and context between and around the variables or parts that make up a system (Gates et al., 2021). For example, a reductionist approach might propose solving a specific issue like unemployment by only creating jobs without considering how education, local industry capabilities, community health, and economic policies would interact with and affect the sustainability of these jobs.

Systems thinking can thus be conceptualised as a “pattern of thought” (Cabrera et al., 2008, p. 306) that can be applied to any discipline or body of knowledge to inform and frame our understanding of the world (Rajagopalan & Midgley, 2015).

While systems thinking emerged in Western academic discourse in the 1940s, it is important to acknowledge that this “pattern of thought” has long been integral to indigenous knowledge systems. Eastern, African and First Nation societies have operated with a systems perspective for millennia, recognising the interconnectedness of social, ecological, and spiritual dimensions of life (Bowman, 2020; Njenga & Massyn, 2024; Pan et al., 2013). For example, African Ubuntu and Ukama philosophy underscores relational interconnectedness, where the well-being of individuals is intrinsically tied to the well-being of their community and environment (Chilisa & Mertens, 2021; Njenga & Massyn, 2024). Recognising these contributions challenges the notion that systems thinking is rooted solely in Western traditions and highlights the need for decolonisation of this field to include diverse epistemologies and worldviews.

However, the remainder of this chapter focuses on Western paradigms of systems thinking in evaluation. This reflects the current dominance of Western academic frameworks in the evaluation field, where much of the literature has been written by Western authors. African and indigenous approaches to evaluation, while vital and growing, are still emerging within mainstream evaluation practice and academic literature. This focus is not intended to overlook or undervalue these contributions but to align with the prevailing body of work that underpins much of the evaluation field today. In Chapter Seven, I seek to address how African approaches to evaluation, in particular, may be integrated into a systems thinking MEL approach.

As a Western method of inquiry, systems thinking encompasses a vast diversity of research approaches in which this paradigm is applied to understand social phenomena. These approaches differ in terms of their underlying ontologies (discussed below) and technical methods. Some authors take a hard view of systems and use computational

methods to understand them (e.g., the field of systems dynamics), while others take a soft view and consider systems from the perspectives of the people who are contained within them (e.g., Soft Systems Methodology; Checkland, 1981).

Let us then first consider what a system is. A system can be defined as a set of interconnected elements or parts that are organised in such a way as to achieve a particular purpose (Arnold & Wade, 2015; Meadows, 2009). This definition can be applied to just about anything, from human bodies to social programmes and the organisations that implement them.

The field's long and expansive history has resulted in a myriad theories, approaches, and modelsⁱⁱⁱ being explored by programme evaluators. Some prominent authors in this space (e.g., Patton, 2011) draw heavily on complexity theory, a branch of systems thinking that defines a complex system as one that adapts to its local environment, is composed of other interacting sub-systems, and is non-linear in nature (Reynolds et al., 2016; Shiell et al., 2008). Important frameworks here are complex adaptive systems (CAS) and living systems theory.

CAS are biological or social systems composed of multiple interconnected components that interact dynamically, learn and adapt to their environment (Holland, 1992). They exhibit emergent behaviour as a result of these interactions, evolving and responding to changing conditions. This means that the whole system's behaviour cannot be fully understood by examining individual parts (Holland, 1992). Examples include ecosystems, economies, and social systems, where adaptability and interconnectedness are crucial to their functioning and resilience.

ⁱⁱⁱ For the sake of brevity, I will use the term 'approaches' as a collective term to refer to systems thinking models, ideas, concepts, methods, methodologies, theories, frameworks, etc.

Living systems theory examines how living systems, from cells to societies, function and sustain life through the interaction of physical matter, energy and information – emphasising how biological and social systems interact to maintain stability (Miller, 1978). This theory bridges biological and social systems, illustrating their interdependence and capacity to process information and energy for adaptation and survival. These concepts underscore the idea that systems are not static but are constantly responding to and influencing their contexts.

Complexity theory and-concepts are pervasive in systems literature, to the extent that systems- and complexity concepts are often inextricable. Some authors simply collapse them into a single term, referring to ‘STCS’ — systems thinking and complexity sciences (e.g., Gates, 2016; Reynolds et al., 2016). Per the terminological note at the introduction to this thesis, I use the term *systems thinking* throughout as an umbrella term, which incorporates complexity theory.

Before taking a critical look at the ways in which systems thinking has been adopted in programme evaluation, it is necessary to briefly review the evolution of systems thinking, to fully contextualise the different approaches and paradigms the field contains.

Systems Thinking Ontologies

Three schools of thought or ontologies have characterised the evolution of systems thinking as a discipline: hard systems, soft systems, and critical systems (Cabrera et al., 2023; Gates et al., 2021; Midgley, 2006). This section follows Gerald Midgley’s (2006) historical account of systems thinking, in which he argues that three definitive waves of research helped to shape the field (discussed further in Cabrera et al., 2023).

The First Wave: Hard Systems

After World War II, three pivotal scientific fields of inquiry emerged: general systems theory, which looks at the organisation of system components; cybernetics, the study of system feedback mechanisms; and complexity science, which explores the behaviours of complex systems (Von Bertalanffy, 1950; Weaver, 1948; Weiner, 1948).

These disciplines introduced a nuanced understanding of causality as something that emerges from interactions between parts within a system, a perspective in sharp contrast to the dominance of reductionism at the time, in which the parts themselves were emphasised (Agazzi, 1978; Cabrera et al., 2023). Scholars at the time took a 'hard' view of systems — that they were real-world entities that could be observed and quantified (Cabrera et al., 2023; Midgley, 2006). As such, the first wave of systems thinking was primarily situated in engineering and operational research, and remains central to various engineering fields today. However, given the universality of systems in the natural world, systems thinking as a philosophical approach also grew at this time - emphasising the common structural elements that define systems across different field or domains (Laszlo, 1972).

By the 1970s and 1980s, however, social scientists began questioning the positivist and realist view of hard systems (Cabrera et al., 2023; Midgley, 2006). Several authors had argued for the use of systems theories in social, and particularly organisational, domains (Emery, 1969).

The Second Wave: Soft Systems

In this period, the notion that systems models should be applied without considering people's role in systems, particularly their perspectives, was challenged. With a particular emphasis on social and organisation contexts, second-wave thinkers highlighted the subjective and socially constructed nature of systems, taking a 'soft' view that elevated the

role of people's perspectives in the interpretation and construction of reality (Ackoff, 1981; Checkland, 1981; Churchman, 1970).

During this era, authors like Ackoff (1981) and Checkland (1981) promoted the idea that the boundaries of a system, previously thought to be as tangible and real as skin on a body, are subjective and open to interpretation. This led to a more inclusive and participatory approach to research, where the input of various stakeholders became integral to defining problems and shaping solutions (Cabrera et al., 2023; Midgley, 2006). A notable methodology to emerge out of this period was Soft Systems Methodology, which championed the use of diverse perspectives to enhance organisational problem-solving (Checkland, 1981; Checkland & Poulter, 2010).

While the second wave contributed a more nuanced understanding of the social dimensions of systems, it was criticised for not addressing the power dynamics inherent in social interaction (Jackson & Keys, 1982; Mingers, 1980).

The Third Wave: Critical Systems

Influenced by critical social theory, critical systems theory emphasised power, conflict, emancipation, and justice within social systems (Cabrera et al., 2023; Midgley, 2006). Third-wave thinkers encouraged a more scrutinising look at how system boundaries are set, as well as the importance of values in these decisions (Flood, 1997; Flood & Jackson, 1991; Ulrich, 1983, 1987).

Moreover, this wave was characterised by a call for methodological pluralism — recognising that the complexity of real-world problems often requires a variety of analytical approaches to understand and address them effectively (Cabrera et al., 2023; Midgley, 2006). Third-wave thinkers thus advocated embracing the rich diversity of systems thinking methods

and approaches, using a contingency approach to match different approaches to specific contexts (Jackson & Keys, 1987; Jackson, 2008).

A Fourth Wave?

Recently, Cabrera et al. (2023) argued that a fourth wave of systems thinking is swelling. While it is unclear whether a full-fledged wave will materialise, the authors argue that there appears to be a discipline-wide shift towards universality and physico-cognitive complexity. As such, this era is reportedly characterised by a search for universal patterns in systems thinking within the physical and knowledge realms, i.e., systems in reality (physical) and how we think about that reality (cognitive).

An important work that the authors draw on to illustrate fourth-wave thinking is that of Cabrera and Colosi (2008), who developed a universal theory of systems thinking: distinctions, systems, relationships, and perspectives (DSRP) theory.

Distinctions (D): Emphasises the importance of identifying what a thing is and what it is not. Making distinctions involves recognising the identity and otherness of any element within any context, thus forming a clearer understanding of each component within a system.

Systems (S): Recognises that every idea or thing can be organised into parts and wholes. It highlights the structure of, and relationships between, parts and wholes, providing insights into how systems function.

Relationships (R): Focuses on the connections between system elements, including the nature, type, and dynamics of these relationships. Understanding relationships is crucial to grasping the dependencies and influences within systems.

Perspectives (P): Recognises that each view or framing of a situation comes from a particular vantage point. This entails understanding that different stakeholders may see problems or systems differently, based on their unique perspectives. Considering multiple

perspectives is crucial to gaining a comprehensive understanding of complex issues, fostering better decision-making and problem-solving.

DSRP theory is particularly significant because it underlines not only the universal patterns of systems thinking in both real-world systems and how we think about them, but it also offers a practical approach to applying systems thinking in everyday scenarios (Cabrera et al., 2015).

Cabrera et al. (2023) argue that systems thinking's potential to survive and thrive in the world today rests on the field's ability to be both approachable (especially concerning newcomers and fresh minds) and mature in its legitimacy. The universal and pragmatic nature of the DSRP theory supports both these goals (Cabrera et al., 2023).

Each wave of systems thinking has advanced the field immeasurably, resulting in a sizeable array of available systems thinking belief systems, methodologies, and research approaches. This overview highlights that systems can be seen as physical realities, social constructs, or arenas of power and ethics, with each perspective bringing its own insights and tools to address the complexities of the social world. While this historical narrative may appear to segregate the systems thinking ontologies into neat time frames, this is not the case. Midgley (2006) and Cabrera et al. (2023) specifically use the term *wave* in their accounts to illustrate that new approaches emerged from a specific time period and context, but that they continue to evolve and overlap with consequent waves. Approaches born out of each wave have not ebbed into obscurity as a new wave appears. First-wave ideas and approaches are still in use today. Rather, the systems thinking field benefits from an ever-increasing expansion of ideas, influencing and shaping each other in turn.

Systems Thinking and Evaluation

Just as the first-wave systems thinkers proposed in the 1940s, there has been a long and growing movement against a reductionist and linear approach to cause-and-effect models in social intervention and programme evaluation. Systems thinking (including complexity theory) has received widespread attention in the evaluation community as a disciplinary alternative to the linear evaluation approaches that strip social programmes of their complex contexts, such as the approaches described in the previous chapter.

Systems thinking is argued to equip evaluators with the tools, concepts, and language to actively engage with complex social situations (Gates et al., 2021; Reynolds et al., 2016). This is not a new development. In the 1980s and early 1990s, authors described the use of systems thinking and complexity methods in evaluation (Gregory & Jackson, 1992; Ulrich, 1988). Today, several books are available on the intersection between the fields (e.g., Hummelbrunner, 2011; Patton, 2011; Williams & Imam, 2006; Wolf-Branigin, 2013). There are a vast number of published case studies using systems approaches to evaluation (e.g., Bustamante et al., 2021; Hassmiller Lich et al., 2017; Osman et al., 2024; Reed et al., 2021; Renmans et al., 2020; Richardson & Patton, 2021), and programme theory (e.g., Chapman et al., 2023; Douthwaite et al., 2020; Douthwaite & Hoffecker, 2017; Wilkinson et al., 2021; Zazueta et al., 2021). The topic has also occupied entire issues of academic journals (e.g., Barbrook-Johnson et al., 2021; Befani et al., 2015; Gates et al., 2021).

Uses of Systems Thinking in Evaluation

Jackson (2022) argues that there have been three major ways in which systems thinking has been incorporated into evaluation. First is the single-methodology approach, in which evaluators draw on a single systems-thinking methodology to conduct an evaluation. From the wave of hard systems thinking, examples include evaluations using system dynamics

(Burke, 2006; Fredericks et al., 2008); from the wave of soft systems thinking, examples include Soft Systems Methodology (Attenborough, 2006; Sgourou et al., 2012); and from the wave of critical thinking, examples include Critical Systems Heuristics (e.g., Gates, 2018; Reynolds, 2006). The disadvantage of single methodologies is the necessity that evaluators be well-versed in many divergent methodologies that vary in suitability for different evaluation situations (Jackson, 2023a; Miller, 2016; Williams, 2019). As Williams (2019, p. 7) puts it: “Learning one method is difficult enough and learning a range is a lifetime’s work”.

Jackson (2022) describes a second approach to systems-based evaluation in response to single methodologies' theoretical burden on evaluators — seeking commonality across the systems thinking field. This is done by identifying the core concepts that characterise different approaches and methodologies across the field’s evolution. First described in an anthology on systems approaches to evaluation by Williams and Imam (2006), these concepts are systems, perspectives, and boundaries. Williams (2019) later described how these concepts were identified: three days of deliberation by evaluation and systems thinking experts.

DSRP theory, described earlier, is another example of a systems concepts approach. While Cabrera and colleagues drew on cognitive science to develop DSRP theory (Cabrera et al., 2008, 2015; Cabrera & Colosi, 2008), another concepts approach draws on the field of Human Systems Dynamics (HSD) (Eoyang, 2006). Similar to DSRP theory, HSD posits its own core set of concepts: containers (boundaries around systems), differences (amongst system parts), and exchanges (between system parts).

The systems concepts were taken forward by the Systems in Evaluation Topical Interest Group (SETIG) of the American Evaluation Association, which published *Principles for effective use of systems thinking in evaluation*, in which the core systems concepts are used as a generic guiding framework for systemic evaluation (SETIG, 2018). Drawing on the core

concepts put forth by Williams and Imam (2006), DSRP theory (Cabrera et al., 2015), and HSD (Eoyang, 2006), the SETIG principles are as follows:

1. Systems-in-evaluation principle: This principle encourages viewing evaluation through a systems lens, incorporating interrelationships, perspectives, boundaries, and dynamics in an integrated manner.
2. Interrelationships principle: This principle focuses on understanding, mapping, and tracking the interrelationships that influence both the evaluand^{iv} and the evaluation.
3. Perspectives principle: This principle highlights the importance of capturing and deliberating on diverse perspectives. Evaluators are encouraged to identify and represent different perspectives and consider the power dynamics associated with each.
4. Boundaries principle: This principle emphasises the need to set and justify boundaries related to both the evaluation and the evaluand. Evaluators are guided to be transparent about their boundary decisions and open to revising them.
5. Dynamics principle: This principle brings attention to the emergent and changing patterns of interactions within systems. Evaluators are encouraged to design evaluations that are responsive to these dynamics, incorporating real-time learning and understanding of complex systems behaviours.

^{iv} An evaluand refers to the intervention, project, policy or programme that is the focus of an evaluation.

The third and final approach to systems-based evaluation discussed by Jackson (2022) is the critical systems thinking approach, described above in systems thinking's third wave of evolution. Drawing on critical systems, Jackson (2022, p. 631) proposes an evaluation methodology that assesses an intervention from five systemic perspectives: the *machine* perspective (efficiency and efficacy), the *organism* perspective (viability and resilience), the *cultural/political* perspective (effectiveness, mutual understanding, and conflict resolution), the *societal/environmental* perspective (marginalised stakeholders, sustainability, and the environment), and the *interrelationships* perspective (leverage points and unintended consequences). This evaluation approach is the fourth step in what Jackson calls 'critical systems practice' (CSP), a four-stage process for intervening in complex situations, informed by the above five systemic perspectives (Jackson, 2020, 2021, 2022, 2023a). The four stages (known as the EPIC Framework) involve: (1) Exploring the situation from the five perspectives; (2) Producing an intervention strategy; (3) Intervening in a flexible and adaptive manner; and (4) Checking, via evaluation, the outcomes and overall effectiveness of the intervention, and refining as necessary.

Critiques and Debates. It is necessary to pause here to acknowledge the ongoing debate between the systems concepts and critical systems thinking approaches to evaluation, publicised in an online debate by the European Evaluation Society (EES) in 2023, between Michael Quinn Patton and Michael Jackson (see Schmidt-Abbey, 2023, for a summary).

The debate centres on the application of systems concepts versus CSP in evaluation, beginning with an article written by Jackson (2023a) describing the evaluation stage of CSP. In this article, Jackson challenges the prevailing use of systems concepts in evaluation, critiquing this approach as "philosophically untenable" (p. 619), having "limited use" (p. 620), and lacking in theoretical rigour and depth. Jackson is not alone in holding these sentiments (see

also Miller, 2016). The criticism prompted a response from Patton (2023), who, in a letter to the editor, defended the systems concepts approach by highlighting its practicality, widespread applicability, and ease of understanding for practitioners in real-world settings. Patton (2023) argued that the concepts approach represents “minimum specifications” (p. 634) sufficient to guide evaluators in thinking systemically. Patton further argued that CSP is too complicated and demanding for most real-world evaluation settings. In the same journal, Jackson (2023b) responded to Patton’s (2023) letter, arguing that CSP is an ideal or a “maximum specification” (p. 637), and that between this maximum and the systems concepts minimum lies a “golden mean” (p. 637) in systemic evaluation usage.

As will be described in Chapter Six, the current research is framed by the systems concepts approach rather than a single methodology or CSP. I agree with Patton’s pragmatic perspectives, and as a newcomer to systems thinking, I can attest to the need for minimum specifications. From my perspective, a major limitation of CSP is that it is not an approachable methodology for someone not well-versed in systems thinking. While Jackson (2023a) acknowledges the “baffling array” (p. 625) of systems thinking approaches and the need to “cut through the confusion” (p. 625), CSP necessitates an in-depth familiarity with several methodologies. CSP’s five systemic perspectives are lenses for viewing complex problems, but are also meant to guide the selection and application of systems methodologies within the EPIC framework (Jackson, 2020). Thus, CSP users are required to appreciate a variety of systems approaches, choose appropriate methodologies, and then select appropriate associated methods (Jackson, 2021). For example, Jackson (2021) lists the following methodological options: systems engineering, the Deming management method, the Vanguard method, socio-technical systems thinking, viable systems diagnosis and design, interactive planning, strategic assumption surfacing and testing, Soft Systems Methodology,

Critical Systems Heuristics, and system dynamics. This reminds us of the shortcomings of the single-methodology approach to systemic evaluation. I agree with Williams's (2019, p. 7) assertion that learning all these approaches "is a lifetime's work".

Barriers to the Use of Systems Thinking in Evaluation

Despite a growing body of literature on systems-based evaluation, there are several barriers to its universal use. Drawing on interviews with evaluation experts, Walton (2016) identified several barriers to incorporating a systems perspective, including resource limitations, increased time- and data requirements, difficulty overcoming dominant approaches (e.g., randomised control trials), managing the expectations of funders and evaluation stakeholders, limited practitioner knowledge and skills, and limitations of available methods and tools. In a similar interview-based study, Gates (2017) argues for the need to promote creative experimentation in systems approaches amongst evaluators, a change in the evaluation commissioning process, and research on using different approaches in practice. These studies highlight notable practical barriers to a systems-based evaluation approach, and I would argue that little has dramatically shifted since they were published.

Reflecting on my experience as an MEL consultant, prior to this thesis, I had never used systems approaches. The barriers identified by Walton (2016) and Gates (2017) help to partially explain why evaluators can be impeded in their interest in a systems approach. However, I argue that an additional three significant conceptual barriers precede this, shaping an evaluator's engagement with the topic in the first place. These are explored below, using my personal reflections as a self-described newcomer to evaluation in the field of systems thinking.

The 'Unapproachability' of Systems Thinking

Earlier, I described the major advantage of and motivation for DSRP theory: a universal approach to systems thinking makes the field more approachable to newcomers and, at the same time, helps to mature and legitimise it through an enhanced understanding of “what makes methodologies systemic” (Cabrera et al., 2023, p. 5).

A significant consequence of the field’s expansive evolution is the sheer volume of approaches available to the uninitiated. Cabrera et al. (2023) argue that the field has become siloed over time as proponents stick to their preferred system camps, resulting in “a lack of a consistent message to newcomers to the field, leading to a feeling of unapproachability” (p. 17). The Jackson/Patton debate described above is a further testament to the lack of consistent messaging. Not only is there no universal welcome message to newcomers, but the uninitiated may, in fact, first learn about systems thinking through stark disagreement and contestation, as evidenced by the EES debate.

An evaluator interested in systems thinking may not know whose camp to join or, really, where to start, which was my experience when I undertook this research. I started with what I thought would be a solid introduction: an anthology of case studies in which systems thinking had been applied to programme evaluation (Williams & Imam, 2006). What I expected to be a straightforward introduction was a confusing baptism by fire into systems thinking. Eleven chapters of the book (Williams & Imam, 2006) described 11 systems approaches, such as system dynamics (Burke, 2006), Soft Systems Methodology (Attenborough, 2006), and Human Systems Dynamics (Eoyang, 2006). The anthology includes pioneering authors in their fields who were conducting impressive and innovative evaluations. However, I struggled to find a clear golden thread through the chapters that clearly elucidated exactly what systems thinking is.

Another seminal text in mainstreaming systems-based evaluation is *Systems concepts in action: A practitioner's toolkit* (Williams & Hummelbrunner, 2011). Williams and Hummelbrunner (2011) provide simple introductions to a number of systems approaches that can be applied to evaluation, including the circumstances in which they are appropriate and steps to undertake them. Despite the easy-to-understand language, the book introduces an overwhelming 19 systems approaches, such as Outcome Mapping, the CDE Model, the Viable System Model, and Critical Systems Heuristics. I thought, 'There is no way I can learn all these approaches well enough to practise them', and sunk into 'buyer's remorse' regarding the topic I had chosen for my thesis.

While this anecdotal account is by no means definitive or empirical, it helps to illustrate just how daunting the system thinking field can be — even for someone with a Master's degree in programme evaluation and many years of practical evaluation experience. Cabrera et al. (2023, p. 16) caution against the field being unapproachable:

Any field requires fresh minds and next generation adoption, and this is very relevant for systems thinking at the present time, given the highly complex global-to-local problems that we are experiencing, where systems approaches are sorely needed.

Wicked Problems

A second significant barrier to using systems thinking in evaluation is the widely held belief among evaluators that these concepts apply only to particular situations, sometimes referred to as 'wicked problems'. Many social problems are perceived as wicked as opposed to 'tame', a distinction introduced by Rittel and Webber in 1973. Tame problems belong in the realm of science, mathematics, and engineering; they are well-defined, and their solutions are easily observable and evaluable as either right or wrong (Rittel & Webber, 1973). On the other hand, the authors include all social policy problems under the banner of 'wicked

problems'. These problems are difficult to define, strongly stakeholder-dependent, and dynamic; they consist of complex, interacting issues that can result in emergent problems as a consequence of trying to solve them (Ritchey, 2013). These problems are not associated with clear, observable, and widely agreed-upon solutions. Rather, solutions are stakeholder-dependent and never complete (Rittel & Webber, 1973).

One of the best-known complexity-aware evaluation approaches is Patton's (2011) developmental evaluation (DE). Drawing on systems and complexity concepts, DE is aimed at evaluators who work collaboratively with social innovators whose programmes are in a state of adaption and/or ongoing development in an uncertain context (Patton, 2011). The evaluator forms part of the programme team, assisting and monitoring and evaluating innovative strategies for programme design as they are implemented.

Patton and others (e.g., Funnel & Rogers, 2011) draw on the Stacey matrix to describe evaluation situations in which complexity-aware approaches, such as DE, are appropriate. In his 1992 book on organisational management, Ralph Stacey described organisations as complex adaptive systems that require a contingency approach to management, depending on the current situation (Stacey, 1992). The matrix consists of two dimensions: agreement and certainty. Certainty refers to how predictable cause-and-effect relationships are, whereas agreement refers to the extent to which stakeholders agree on the way forward. Stacey (1992) argued that conventional linear approaches to decision-making are inadequate in situations of high uncertainty and disagreement (i.e., complex situations).

A similar heuristic used by evaluators is the Cynefin framework, which also categorises contexts in terms of their complexity and predictability (Snowden, 2005; Snowden & Boone, 2007). While the Stacey matrix primarily maps situations based on the levels of certainty and agreement, Cynefin considers the nature of cause-and-effect relationships to guide decision-

making. The framework map contexts into four domains, each with distinct cause-and-effect relationships and thus requiring different decision-making strategies: clear; complicated; complex; and chaotic.

The Stacey matrix and Cynefin framework has been adopted as situation recognition heuristics by some evaluation authors (e.g., Funnel & Rogers, 2011; Patton, 2011; Westley et al., 2011; Zimmerman et al., 1998) to distinguish the simplicity versus complexity of social situations. Simple, or clear, situations have high degrees of certainty and agreement, and straight forward cause-and-effect relationships, thus requiring equally simple interventions (Patton [2011] used the example of an immunisation campaign), whereas complex situations have high uncertainty and disagreement, and unpredictable causal relationships. Authors such as Funnel and Rogers (2011) and Patton (2011) used these heuristics to characterise the complexity of an evaluand's situation. The more complex a situation is, the more appropriate a complexity/systems approach will be to its evaluation (e.g., DE). As such, Rogers (2011) and Patton (2011) propose a contingency approach to evaluation, whereby systems thinking is only relevant in particular contexts; contexts that exemplify high uncertainty and disagreement. At face value, this could suggest that the application of systems thinking to evaluation is not inherently useful or valuable in all contexts, but only rather in overtly complex contexts according to heuristics such as the Cynefin framework of Stacey matrix.

Mowles (2014), a prolific writer on complexity in management, argues that some authors show a problematic tendency towards overstating the contingency approach in evaluation. Mowles (2014) argues that complexity should be taken as the default; social life is inherently emergent, irrespective of whether a programme is tightly or loosely planned and organised. Patton's (2011) aforementioned example of an immunisation campaign as a simple

situation (with high levels of agreement and certainty) is the perfect example of the naïveté of this perspective.

The global response to the rollout of COVID-19 vaccines proved that an immunisation campaign is far from simple. While there are standards, best practices, and clear cause-and-effect variables in the development of vaccines, the rollout and uptake of COVID-19 vaccines were significantly influenced by complex social, cultural, and informational factors, such as inconsistent government communication and policy (Moss et al., 2022), cultural attitudes towards vaccines (Luo et al., 2021), and misinformation and conspiracy theories, which undermined public perception of vaccine safety and efficacy (Jennings et al., 2021).

In a review of system and complexity concepts in evaluation, Gates (2016) argues that systems concepts have value for *any* evaluation purpose and intervention context, including the design, implementation, and management of social interventions. Further, Walton et al. (2021) reject the narrow view of systems-based evaluation as a means to value a programme or process, and argue for an expanded application, such as framing problems, constructing strategic models for change efforts, and learning and adaptation. I agree with this position, and believe that systems thinking has far more to offer both evaluation processes and programme design than is generally applied. Systems thinking, as a discipline, is not positioned as only appropriate in particular situations, and it should therefore not be perceived as only applicable to particular evaluands (working in wicked spaces).

This is especially true in the context of social interventions as delivered by NPOs in developing countries like South Africa. Local NPOs, arguably, face complexity both in terms of operational management and designing programmes to tackle complex social problems. I illustrate this in more detail in the following chapter, which discusses the complexity of early childhood development. I would go so far as to argue that there is only complexity where

social intervention is concerned. Human beings are themselves complex creatures, operating from within layers of subjective perception and interacting with, and themselves influencing, a multitude of social systems (e.g., the family system, the school system, neighbourhoods, governments, social circles, etc.).

In my view, evaluators' use of the Stacey matrix and similar heuristics to situate systems-based evaluation tools or approaches (such as DE) in complex situations may inadvertently lead evaluators to think of systems approaches as suitable only for complex or wicked situations. This could discourage evaluators working with 'simple' programmes from considering systems approaches, despite the many advantages to doing so, as illustrated throughout this thesis.

Overemphasis on Summative Evaluation

A final barrier to entry into the systems thinking field concerns the overemphasis on summative evaluation. While the evaluation field contains several approaches for the formative tracking of programmes in uncertain contexts, such as Real Time Evaluation (RTE) (e.g., Polastro, 2014), evaluation scholars wishing to incorporate systems thinking concepts have primarily applied these concepts to summative evaluation.

With the exception of DE, which is known as a pre-formative approach to evaluation (Patton, 2011), the majority of other available evaluation approaches apply systems and complexity concepts to the design of summative evaluations, guiding external evaluators who are conducting outcome or impact evaluations of programmes deemed to be complex, or programmes aimed at changing systems (e.g., Hargreaves, 2010; Latham, 2014; Parsons, 2007; Preskill et al., 2014; Renger, 2015; Urban et al., 2021; Wasserman, 2010). This is arguably due to the contingency framework discussed above, in which complex programmes or complex systems change initiatives are deemed suitable for such evaluation designs. There

is a significant lack of comprehensive guidance on using systems thinking in MEL in generic programmes, that is, programmes that might not be deemed ‘complex’ using the Stacey matrix. Four exceptions are described below.

In the aforementioned book, *Systems concepts in action: A practitioner’s toolkit*, Williams and Hummelbrunner (2011) describe an approach they call ‘Process Monitoring of Impacts’, which was developed by Hummelbrunner to support the monitoring practices of European Union Structural Fund programmes. The emphasis in this approach is on the process of generating results, rather than just “indicators of their final measure” (Williams & Hummelbrunner, 2011, p. 93). The problem with their approach is that it is simply inaccessible. Beyond the information provided in the book’s chapter, I could not find any resources to guide its application.

In 2020, USAID developed a guidebook on complexity-aware approaches to the monitoring of programmes that it funds within the MOMENTUM portfolio (aimed at maternal and child health) (USAID, 2020). Similar to the evaluation toolkit of Williams and Hummelbrunner (2011), the guidebook describes several systems approaches that can be selected and applied to programme monitoring, such as social network analysis, Outcome Harvesting, and contribution analysis. In my view, this resource suffers the same approachability limitation — it is difficult to determine which approach to select from the list and how to apply it.

Reflexive Monitoring in Action (RMA) was developed by Van Mierlo (2010) for programmes in system innovation for sustainable development. RMA is similar to, but far less known than, Patton’s (2011) DE, in that it was specifically designed as an integral part of adaptive programme design amongst a heterogenous group of developers. Drawing on adaptive programme management, RMA is a guiding framework aimed at helping innovative

programmes respond and adapt to their dynamic contexts to meet real-time emerging challenges and opportunities. As with DE, RMA is a contingency approach, designed for use in a specific pool of innovative programmes.

Finally, one of the most recognised systems approaches that can assist evaluators and organisations with the development of a MEL framework is Outcome Mapping (OM). OM was developed in 2001 by Earl et al. for the International Development Research Centre.

OM comprises practical steps that a programme team takes, along with an evaluation facilitator, to map the programme's objectives, strategies, and M&E plans (Earl et al., 2001). OM is a highly participative approach to MEL, using a workshop format with programme staff. OM was designed for use in programmes that do not have direct influence over their ultimate intended beneficiaries but rather work with boundary partners — intermediary actors situated between the programme and its ultimate intended beneficiaries. The approach is based on the premise that change within a system is created by changing people's behaviour, particularly in terms of relationships and collaboration, within the programme's specific sphere of influence (Outcome Mapping Learning Community [OMLC], 2021).

OM is a fairly radical departure from traditional MEL through its emphasis on qualitative programme monitoring and reliance on its own jargon. For example, instead of using indicators, OM proposes the use of 'progress markers'. Progress is tracked using journals of outcomes, strategy, and organisational performance. While these elements are not inherent weaknesses of the approach, they do limit its utility and feasibility across settings, particularly if programme managers are required by funders to report on, for example, particular quantitative indicators. The OM jargon was noted as a particular barrier in a 10-year retrospective study of the approach's utilisation (Smith et al., 2012). It has also been criticised for being complex and both time- and resource-intensive to set up and maintain in

organisations, due to its reliance on sensemaking amongst boundary partners with regard to whether change has taken place, as well as its emphasis on behavioural outcome data (Jones, 2006; Smith et al., 2012). Sensemaking is an important part of OM, because it is based on the assumption that change is emergent and non-linear, and thus cannot be predicted (employing CAS perspective) (OMLC, 2021). OM therefore seems to be another contingency-based approach that is only suitable for programmes whose outcomes cannot be predicted with certainty and agreement. As with DE and RMA, OM is intended to be a highly adaptive programme management approach aimed at balancing goal-setting and accountability with “experimentation, probing, and discovery” (OMLC, 2021, p. 8).

In closing, the use of systems thinking in evaluation appears to be skewed towards summative or external evaluations of programmes. While there are some promising approaches to MEL in particular, these have been designed to focus on wicked problems or complex systems change, and are thus only suitable for organisations that have the capacity for adaptive programme management and the freedom to implement programmes with high levels of flexibility and experimentation. In my experience, such autonomy and operational flexibility are rare in South African NPOs, who are typically constrained in terms of organisational capacity and accountability reporting to funders. Additionally, in my view, MEL approaches based in the CAS approach are not necessarily appropriate in all contexts of social change. Sometimes, even in complex social settings, change can be predicted using existing empirical evidence, social science theory, and the experiential insights that NPOs develop over time. Based on the above, MEL could be more efficient using a predictive model of change, rather than wholly relying on tracking emergent processes. This argument is explored further in the following chapter, specifically in the context of ECD.

Chapter Conclusion

This chapter described the depth and utility of systems thinking as a critical theoretical framework for MEL in social programmes. The chapter described how systems thinking evolved in response to the limitations of scientific reductionism. This evolution was framed according to three significant waves of systems thinking — hard, soft, and critical — each reflecting shifts in focus from the tangible aspects of systems to more interpretative and critical perspectives that consider power dynamics and social constructs. Each wave has contributed additional insights and perspectives that are still relevant in understanding social phenomena and, thus, social interventions.

The application of systems thinking in programme evaluation is advocated as a means to transcend traditional linear and reductionist approaches, which often fail to capture the nuanced dynamics of social phenomena and change. Rather, systems thinking (and all its waves) offers a holistic view that more accurately reflects the systemic nature of social phenomena. This chapter, however, also presented a number of substantial barriers to the adoption of systems thinking in MEL, including the 'unapproachability' of systems thinking, the contingency approach to applying systems thinking, and an emphasis on summative evaluation. Each of these barriers represents conceptual and practical hurdles that evaluators face, from the overwhelming diversity of systems thinking approaches, which may deter newcomers, to the belief that systems thinking is only applicable in specific (and often complex) situations, which could limit its use in everyday settings.

In conclusion, this chapter argues for the adoption of systems thinking in MEL, suggesting that a shift towards this approach is not merely beneficial, but essential. However, the evaluation field needs additional resources for systems thinking newcomers who consider

themselves to be working in routine or everyday settings and are unfamiliar with the systems thinking territory.

Chapter Four: ECD as an Interesting Test Case

A caregiver who is burdened by poverty and domestic violence, and is very distressed, may struggle to provide the quality of care she desires to provide for her infant. A consequence could be malnutrition, which affects neurological development, in turn compromising cognitive functioning and skills development at school, as well as undermining emotional wellbeing and a secure sense of self. In adolescence, failing to progress in school together with the rising costs of education may lead to intermittent attendance and early departure, dashing individual and family ambitions for a better future. (Boyden et al., 2019, p. 147)

In the extract above, Boyden et al. (2019) paint a somewhat disheartening picture of the complex consequences of childhood poverty on human development, highlighting the multiple different yet interacting dimensions that shape the life course.

As described in the introduction, I primarily conduct MEL and evaluation consulting with ECD clients. I know this sector well and continuously strive to participate meaningfully in it. Given my experience and networks in the field, I decided to research the use of systems thinking in MEL in an ECD setting. Further, I believe that ECD is a particularly interesting test case for a systems thinking approach to MEL, which I explore in this chapter.

This chapter delves into the complexities of ECD as portrayed in Boyden et al.'s extract above. Using South African and global literature, I explore how child development is shaped by various interconnected systems, particularly in resource-deprived contexts such as South Africa, using Bronfenbrenner's (1979) Socio-ecological Model. While acknowledging the complexities of ECD, this chapter also emphasises what we *do* know and what we can predict, arguing that ECD is perhaps an interesting test case that delegitimises the Stacey matrix and contingency approach to systems-based evaluation (discussed in the previous chapter).

The Importance of ECD

ECD is increasingly recognised as critical to the health and development of people and societies in both developed and developing countries (Britto et al., 2018). Growing international evidence continues to demonstrate that adverse experiences in an individual's early years have long-term consequences for their life trajectories (Black et al., 2017). However, early intervention can prevent developmental delay and improve longitudinal adult outcomes such as employment and health (Anderson et al., 2003; Black et al., 2017). ECD programmes are designed to intervene in these early years and improve the physical, cognitive, and socio-emotional functioning of children in preparation for formal schooling (Anderson et al., 2003).

The advantages of promoting ECD have received global recognition and political commitment, evident in its inclusion in the United Nations Sustainable Development Goals (SDGs) (United Nations, 2015). SDG 4.2 states that, by 2030, countries should “ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education”. Aligning with this global initiative, the South African government introduced several policy initiatives, including the Education White Paper 5 of 2001, the National Integrated Plan for ECD (2005–2010), and the National Integrated ECD Policy of 2015. In all three, the provision of quality ECD services is prioritised, including the need for universal and equitable access. The need for equitable access is also reflected in the National Development Plan, which includes a section on ECD, specifying the need to expand ECD programmes to all vulnerable children, and recommending that all children access two years of ECD services before entering Grade 1.

Despite significant advancements in South Africa's ECD policy framework, there are a number of factors that hinder effective policy implementation. Two local studies were

conducted to map South Africa’s funding and programmatic landscape for ECD and reveal important gaps in the sector (Horler & Biersteker, 2022; Horler et al., 2023).

The Complexity of ECD

Child development is perhaps conventionally thought of as a linear process — a steadily increasing line on a growth chart. However, children do not develop in a vacuum, particularly in resource-deprived contexts which are pervasive across South Africa.

Approximately 70% of children in South Africa (aged 0 to 17 years) live in households below the upper-bound poverty line^v, indicating that these families struggle to meet both essential food and non-food needs (e.g., shelter and clothing). The adverse impact of poverty on human development is well established. Longitudinal studies have found that childhood poverty can have a detrimental effect on academic achievement, behavioural outcomes, health status, earning potential, and socio-economic well-being (Boyden et al., 2019; Duncan et al., 2011).

In the sections that follow, I draw on the Socio-ecological Model of child development (Bronfenbrenner, 1979) to illustrate the complexity of child development, particularly in contexts of poverty, such as those prevalent in South Africa.

The Socio-ecological Model

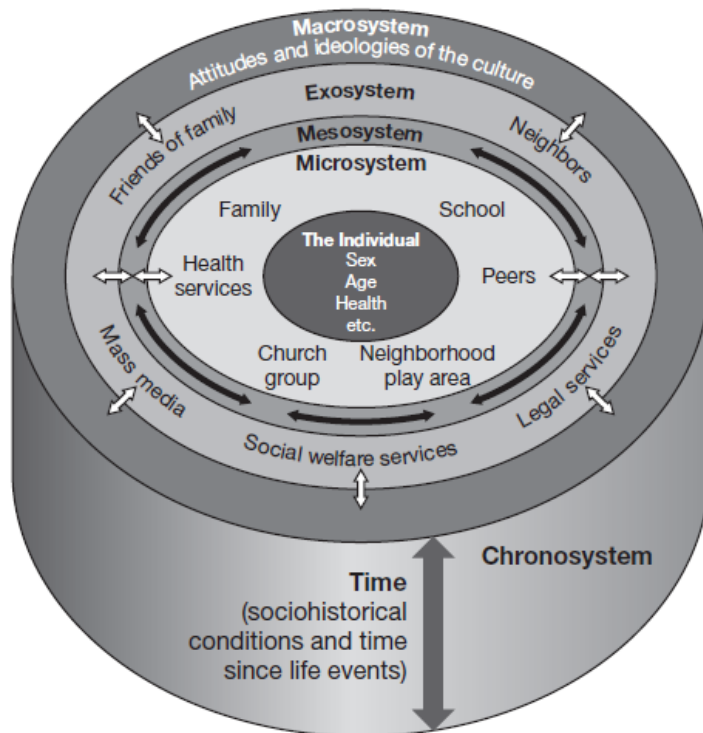
Bronfenbrenner’s (1979) socio-ecological model (SEM) of child development is extensively used to understand the nested, interacting systems of development, which

^v The upper-bound poverty line in South Africa, as defined by Statistics South Africa (2022), is an income level at which households have just enough resources to cover both essential food and non-food needs, such as clothing, transportation, and education.

include: individual children and their attributes; children’s microsystem with which they interact directly (e.g., family, peers, school, church); the mesosystem in which microsystems interact with each other (e.g., parents + school; peers + church; family + clinic); the exosystem which contains influences in the broader environment (e.g., neighbours, media, NPOs, social services); the macrosystem, which contains the cultural and societal influences over all preceding systems (e.g., politics, economic conditions, ideology, cultural and social norms, policies); and, finally, the chronosystem, which contains historical events and changes that occur over the life course (e.g., the COVID-19 pandemic, economic depression). As illustrated by the arrows in Figure 2, all the socio-ecological systems interact with and influence each other.

Figure 2

Bronfenbrenner's Socio-ecological Model



Note: From Introducing Bronfenbrenner: A Guide for Practitioners and Students in Early Years Education (p. 14), N. Hayes, L. O’Toole, & A. Halpenny, 2017, Routledge. Copyright 2017 by N. Hayes, L. O’Toole and A. Halpenny.

The SEM (Bronfenbrenner, 1979) illustrates how children’s growth and developmental outcomes are not achievable through only one system or pathway. The model shows that, instead, there are multiple levels of influence situated in different but interacting systems. In the paragraphs that follow, I use the layers or systems indicated in the SEM (Bronfenbrenner, 1979) to describe the complexity inherent in South Africa’s ECD sector.

Macro- and Exosystems. Government policy, which reflects the macrosystem, has a substantial influence over children’s outcomes and long-term trajectories, as it dictates the rights and services to which children should have access. South Africa’s National Integrated Early Childhood Development Policy (NIECDP), approved by cabinet in 2015, provides a comprehensive framework of a broad range of essential services that must be made available to all children, focusing especially on uplifting South Africa’s poorest children. These services include health and nutrition, social protection, parent support, opportunities for learning, and services related to an adequate standard of living, e.g., water and sanitation (Republic of South Africa [RSA], 2015). The policy recognises the need to foster inter-system collaboration between the government and the non-profit sector to integrate the provision of these services.

Historically, non-profits in South Africa have provided the bulk of ECD services to communities, often developing innovative programme models that get adopted by the government, and training an ECD workforce to provide services to children (Atmore et al., 2012). Training and resource provision remains a key part of ECD NPOs’ offerings to the sector, particularly in the area of early learning; 89% of NPOs operating in early learning report that training and resource provision are their core offerings (Horler et al., 2023).

According to Hall et al. (2019), little progress has been made by the government regarding the integration of services, due to fiscal constraints and “inadequate institutional

mechanisms required to co-ordinate, manage and monitor complex integrated ECD service delivery” (p. 7). The majority of NPOs in the sector are primarily funded by donors, with decreasing financial support from the government (Horler et al., 2023). This means that NPOs often design and operate their programmes according to donor priorities to acquire funding.

Donors therefore play a powerful role in the sector, often determining how the NIECD is implemented. An example of this is the way early learning programmes (ELPs), such as formal and informal preschools, crèches, or playgroups, are subsidised. While the NIECDP calls for an increase in equitable access to early learning services for poor children, ELPs are not fully subsidised by the government in the way that primary and secondary education institutions are (Hall et al., 2019). This means that ELPs need to source their own funding, and they rely primarily on school fees (Department of Basic Education [DBE], 2022). In communities where parents and primary caregivers live in extreme poverty, this has two critical implications: children do not attend ELPs because their families cannot afford the fees (Hall et al., 2019), and ELPs do not generate sufficient funds to sustain quality services for children (Biersteker et al., 2016).

ELPs can access government funding by achieving formal registration and compliance with specific standards. To qualify for government subsidies, these programmes must meet health, safety, and educational standards as outlined in the Children's Act and comply with a number of norms and standards that cover aspects like infrastructure, staff qualifications, and programme content. To help ELPs unlock government funding, many NPOs focus their efforts on supporting ELPs to meet these standards by upgrading their facilities into more formal ECD centres (Horler et al., 2023). This reflects a key donor priority: funding ECD centres is the most common investment area among both South African and global donors (Horler & Biersteker, 2022). Therefore, donors align with the government’s priority of ELP formalisation

and registration, and thus fund NPOs to support this. More informal ELPs, such as home-based playgroups, often cannot satisfy the criteria for government registration and thus cannot access subsidies. As these programmes are less aligned to government priorities, they receive less attention by donors and NPOs. Less than half of global funders report investing in playgroups compared to ECD centres, and about 10% of NPOs that report practitioners from ECD centres report training playgroup practitioners (Horler et al., 2023).

This examples reveals the critical interactions between stakeholders in the macro- and exosystems (government, donors, and NPOs) that shape access to services in other systems in which children and their families are located. While the NIECDP aspires to equitable and integrated service delivery, its implementation often reflects the power dynamics of funders, who prioritise formalised programmes that align with government registration criteria. This emphasis on formal ECD centres risks marginalising informal programmes like playgroups, which are crucial cost-effective services for reaching underserved communities (Desmond et al., 2019).

The Chronosystem. Implementation of national policy in the macrosystem and exosystem is, of course, influenced by socio-historical events, such as the COVID-19 pandemic, the negative effects of which filter down to all subsequent socio-ecological systems. The ECD sector was severely disrupted by the pandemic and lockdowns, as the government forced ELPs to close their doors and ordered that children be kept home. The effects of the lockdowns were long-lasting. Following a court case, the sector was permitted to re-open, but a year later, child attendance at ELPs had still only partially recovered to pre-pandemic levels (Wills & Kika-Mistry, 2021). ELPs lost their sources of funding, as parents stopped paying fees, and the government was slow and ineffectual in providing financial relief (Bayat & Madyibi, 2022).

The effects of the pandemic on the individual child are wide-ranging. A global scoping study (Penna et al., 2023) noted that the pandemic had caused multiple interacting stressors in the lives of young children, including the death of a parent, greater financial stress on households, loss of access to health services, and decreased maternal mental health (particularly as childcare services became unavailable) (Penna et al., 2023). Other international studies (Bem-Haja et al., 2022; González et al., 2022) have noted severe disruptions in young children's developmental competencies, particularly amongst those living in poor socio-economic conditions.

Micro- and Mesosystems. The availability of community-level services in the exosystem (as determined by the macro- and chronosystems) inevitably affects the family system, which, in turn, directly affects child development. The quality of the home learning environment has consistently been found to influence child outcomes (Connor et al., 2005; Hamadani et al., 2010; Visser et al., 2015). Higher levels of education amongst South African mothers have even been found to buffer the long-term negative effects of early developmental delay (Hsiao & Richter, 2014), as mothers with increased levels of education have been shown to provide more early learning materials and demonstrate increased levels of responsiveness to their children (Magnuson et al., 2009).

Nutritious food, which plays an important role in child development, has also been linked to parental education. A recent study in Limpopo found that many caregivers in receipt of the Child Support Grant had never been taught about child nutrition (Kekana et al., 2020). Less than 50% of the grant was spent on food, and most of the food had low nutritional value. The prevalence of wasting and stunting amongst children in this community was very high, alongside high rates of obesity in caregivers.

These examples illustrate the profoundly complex nature of the systems involved in child development. While child development competencies, such as fine motor skills and emergent literacy, are well understood in terms of children's growth trajectories, the degree of change in these outcomes depends on a complex array of interacting systems at the level of government, communities, and the family.

Thus, it is a major concern that traditional and linear approaches to monitoring and evaluation often fail to take these interacting contextual influences into account, therefore preventing a complete understanding of the child outcomes expected to result from an ECD programme. Systems thinking, discussed below, might be one area that can address these limitations.

The Knowns of ECD

Child-rearing is often used to illustrate maximum complexity on the Stacey matrix. For example, Patton (2011, p. 91) says:

Parenting is complex. Unlike the metaphor of a cooking recipe for a simple situation or the rocket-launching metaphor for a complicated situation, parenting involves huge uncertainties and no clear rules guaranteeing success to follow. Oh, to be sure, there are many parenting experts and many guides available to parents. But none can be treated like a cookbook for a cake, or a set of formulae to send a rocket to the Moon. In the case of the cake and the rocket, for the most part, we were intervening with inanimate objects. The flour does not suddenly decide to change its mind and gravity can be counted on to be consistent, too. On the other hand, children, as we all know, have minds of their own. Hence our interventions are always in relationship with them.

Undoubtedly, we cannot predict the unique qualities, characteristics, and abilities each young child will develop and exhibit. However, in the context of ECD, several known

factors make child outcomes a lot less unpredictable than the Stacey matrix would have us assume. The literature presented above indicates, for example, the sheer breadth of what we *do* know about the systemic factors that influence the development of children's outcomes. As noted earlier, ECD is growing as a global priority, as is the body of literature supporting evidence-based programming.

In terms of young children's ability to learn and develop readiness for formal schooling, large-scale longitudinal studies have identified a number of key cognitive competencies that predict later educational outcomes: emergent mathematics skills, early reading ability, ability to pay attention, fine and gross motor skills, as well as social and emotional skills (Duncan et al., 2007; Ricciardi et al., 2021).

Literature on early learning has led South African researchers to develop a culturally fair instrument to assess young children's developmental skills, available in all 11 of South Africa's official languages. The Early Learning Outcomes Measure (ELOM) (Snelling et al., 2019b) is a standardised instrument that assesses children aged four to five years in terms of their motor, cognitive, language, mathematics, and socio-emotional skills. The instrument also includes an optional growth status measure using height-for-age z-scores that determine whether children are stunted. There is a vast base of evidence supporting the importance of these domains, including growth status, in children's long-term development trajectories (e.g., Arnold et al., 2012; Boyden et al., 2019; Collie et al., 2019; Duncan et al., 2007; Fitzpatrick et al., 2014; Georgiadis & Penny, 2017; Nayfeld et al., 2013; Ricciardi et al., 2021; Siegler et al., 2012).

The ELOM (Snelling et al., 2019b) has established standards for development amongst South African children from diverse socio-economic and ethnolinguistic backgrounds, allowing researchers and evaluators to establish whether children are 'on track', 'falling

behind', or 'falling far behind' these standards (Snelling et al., 2019a). The tool has been used in several national research studies and programme evaluations that assess the effectiveness of ELPs, and has helped to track children's outcomes and determine necessary programmatic inputs (Dawes et al., 2020, 2021; Horler et al., 2019; Tredoux et al., 2024). For example, studies using the ELOM have uncovered significant predictive relationships between ELOM results and ELP participation, teacher–child ratios, stimulation in the child's home environment, growth status, and children's multi-year participation in ELPs (Dawes et al., 2020, 2021; Horler et al., 2019).

In summary, while child development is inarguably a complex process involving the interaction of variables across several dynamic, mutually influential human systems, there is a significant empirical evidence base that has uncovered many agreed-upon 'knowns'. The relationship between what we know and what we cannot know means that "outcomes of early experience are probabilistic rather than deterministic" (Boyden et al., 2019, p. 25). Thus, while we cannot determine definitive causal relationships between programmes and outcomes, research innovations such as the ELOM support a growing local literature that enables researchers and evaluators to more reliably assess the important inputs for successful programmes for early learning. What we cannot predict is the unique variability in child outcomes — why some children may reach milestones quicker than others, or why some children may develop skills in different sequences (Vereijken, 2010). I propose that variability in child outcomes is inherently complex; it is non-linear, emergent, and unpredictable. However, the general domains of these outcomes, such as cognitive or motor skills, and the inputs required for developing them, are well-established and show predictable patterns at a population level.

Chapter Conclusion

Patton (2011) and others use child development as an example of complexity because there are no standard protocols that result in a defined set of outcomes in this sphere. Rather, as explained by the SEM (Bronfenbrenner, 1979), children develop in the context of dynamic, interactive, and multidimensional systems. As such, using the Stacey matrix would necessitate ECD programmes taking an adaptive and experimental approach to programming and MEL, exploring multiple different pathways to child outcomes.

However, the available literature on ECD indicates that an adaptive approach to MEL may be impracticable when considering the time and resources needed to implement, for example, resource-intensive DE, RMA or OM. This chapter argued, based on empirical literature, that ECD presents a compelling alternative to the Stacey matrix by demonstrating how a complex issue like human development can simultaneously embody high levels of certainty and agreement.

ECD reveals the limitations of the Stacey matrix by highlighting how, despite the multifaceted nature of child development, numerous established factors and probable patterns exist. These include the effect of early interventions on longitudinal outcomes and the predictive power of key cognitive competencies with regard to future educational success. The chapter also discussed the efficacy of standardised assessment tools like the ELOM. Such instruments provide consistent metrics with which to evaluate child development across diverse socio-economic and cultural backgrounds, illustrating how standardised approaches can be applied in a complex system.

ECD programme developers can leverage these knowns to design effective programmes and structured MEL systems that use empirical evidence as a basis, instead of investing significant resources and capacity in 'reinventing the wheel' through exploratory and

experimental approaches to programme management and MEL. Systems thinking can help to enhance this process by affording programme teams the language and tools to deliberately address the systemic, multi-layered interactions within the ECD ecosystem and develop an appreciation for the complex factors that help shape the knowns.

Problem Statement and Research Questions

In Chapter Two, I examined the adoption of MEL in South Africa's non-profit sector and prevailing trends in MEL approaches, such as RBM and linear models of programme theory (i.e., logframes and logic models). I argued that these approaches do not support the organisational embeddedness and learning that is central to MEL, due to an overstatement of accountability and a failure to capture the complexity inherent in social interventions.

In Chapter Three, I described the field of systems thinking and the ontological waves that have characterised the evolution of the field over time. I presented literature on the use of systems thinking in evaluation, and highlighted the conceptual benefits that systems thinking can offer the evaluation discipline, most notably in terms of a holistic understanding of social phenomena. Despite its utility, I argued that there are a number of barriers to adopting systems thinking practice amongst evaluators, including the field's perceived unapproachability, a perspective that dichotomises programmes or situations as either complex or not, and a paucity of systems thinking-informed MEL approaches for evaluands that lack capacity for completely adaptive and emergent approaches to programme management and MEL.

In Chapter Four, I expanded my argument that systems thinking is deemed only suitable for supposedly complex situations or programmes, using the field of ECD as an example. Despite the immense complexity that characterises child development, there is solid empirical evidence illustrating what we know and what we agree upon with regard to expected child outcomes. While we cannot predict the variability in child outcomes, or determine causal relationships, we can draw on a vast body of literature to help predict the likelihood of outcomes based on particular programmatic inputs. This provides an efficient and theoretically sound basis for MEL in ECD programmes. However, evaluators lack

guidelines for establishing MEL frameworks that leverage the benefits of systems thinking while also allowing for the knowns of the sector (ECD in this case). As such, the problem statement that underlies this research is as follows:

All social programmes are implemented in complex social contexts, and would thus benefit from MEL structures that are less rigid, linear, and inflexible than the RBM-orientated tools that proliferate today. Systems thinking is well suited to MEL, but there is a lack of approachable guidelines or tools for evaluators who are (a) new to systems thinking and (b) work in social sectors that are characterised by both complexity, on the one hand, and agreement and certainty, on the other, with regard to outcomes.

This research attempts to fill the gap in the practical intersection between MEL and systems thinking through the design and field testing of a systems thinking approach to MEL in ECD. As such, this research was aimed at answering the following research question and sub-questions:

(1) How feasible is a systems thinking approach to MEL development?

(1.1.) How did the workshop participants and the evaluators (workshop facilitators) experience the Systems Approach to MEL Toolkit as a systems-thinking approach to MEL?

(1.2.) What are the challenges in and advantages of integrating systems thinking into MEL?

The following chapter describes the research methodology employed to answer these questions.

Chapter Five: Methodology

As noted in Chapter One, this research had two primary objectives: 1) to develop a MEL toolkit based on systems thinking that can guide MEL framework development in a participatory workshop format, and 2) to test this toolkit amongst South African ECD NPOs.

Toolkit Rationale

I decided to develop a systems thinking-based MEL toolkit due to the lack of user-friendly and accessible guidelines available to evaluators who want to apply systems thinking to MEL. I specifically chose a toolkit as the knowledge product because I intended to create clear, step-by-step guidelines, together with practical worksheets, to assist evaluators in their comprehension of systems thinking.

The toolkit was given the working title: *Systems Approach to Monitoring, Evaluation and Learning* (SAMEL). Its development and content are described fully in Chapter Six. In this chapter, I describe the research design and research paradigm that were chosen to achieve the objectives above, as well as my approach to sampling, data collection, and analyses, together with the rationale for the methodological choices.

Research Phases

This research necessitated multiple steps to design and test the SAMEL Toolkit. Given the sequential nature of the research objectives, the research was divided into two main phases: (1) design and (2) test, followed by a report of the findings. These phases are similar to those of Design Science Research in information sciences. Design Science Research is a problem-solving, pragmatic paradigm in which information science innovations are created (and evaluated) to address complex real-world problems (Hevner & Chatterjee, 2010). Peffers et al.'s (2006) Process Model for Design Sciences Research is as follows: (1) identify the problem; (2) define objectives of a solution; (3) design the solution; (4) demonstrate proof-of-

concept in test cases; (5) evaluate its performance; and (6) communicate the findings. All six steps were incorporated in the current study, but were collapsed into two main phases for brevity (see Table 2).

Table 2

Research Phases Aligned to Design Science Research Steps

Design step	Research phase	Research method	Corresponding thesis chapter
1. Problem identification	Proposal	Literature review	Chapters One, Two, Three, and Four
2. Objectives	Phase one: Toolkit design and expert review	Toolkit development using theory-based evaluation and systems thinking as conceptual frameworks; Expert review of Toolkit prototype	Chapter Six
3. Design			
4. Demonstration	Phase two: Testing the Toolkit	Testing the Toolkit using two NPO cases; Qualitative descriptive and reflexive thematic analyses of the gathered data	Chapters Eight and Nine
5. Evaluation			
6. Communication	Reporting	Reporting and discussion of findings of thematic analysis	Chapter Ten

To achieve each of the research phases, the study adhered to a pragmatist paradigm.

Research Paradigm

A paradigm is the theoretical orientations that act as the lens through which qualitative researchers employ methods, interpret the data, and frame their findings (Creswell & Poth, 2018). A paradigm reflects the researcher’s beliefs about the nature of reality, the nature of knowledge, the role of their values and biases in the research, and the process and nature of data collection and analysis (Creswell & Poth, 2018; Lincoln et al., 2018). Declaring these beliefs assists the reader to understand the design decisions made by the researcher (Creswell & Poth, 2018; Maxwell, 2013).

The worldview governing the present research was pragmatism. Pragmatism is primarily concerned with providing practical solutions and useful applications to concrete real-world issues or problems (Patton, 2015b). Pragmatism is rooted in the work of early 20th

century philosophers Dewey, Mead, James, and Pierce, who argued that knowledge is generated through action, and that the meaning of an idea lies in its practical consequences (Corbin & Strauss, 2008; Goldkuhl, 2012; Patton, 2015b). Pragmatic inquiry is thus driven by the purpose of going beyond observation, to create knowledge that will inform the desired change or improvement (Goldkuhl, 2012). Goldkuhl (2012) calls this ‘constructive knowledge’ — knowledge that provides understanding and is descriptive, but is also prescriptive (providing guidelines), normative (value-driven) and prospective (suggesting possibilities). The SAMEL Toolkit is arguably a constructive knowledge product.

Thus, a pragmatist inquiry is guided by the intended practical consequences of the research and, as such, employs a variety of methods based on how well they answer the research questions (Creswell & Poth, 2018; Creswell & Creswell, 2018). Pragmatist research remains flexible and open; there is no strict allegiance to a particular tradition, method, or philosophical worldview (Creswell & Poth, 2018). Table 3 summarises the pragmatic beliefs that underpinned the present research.

Table 3*Pragmatic Beliefs that Underpinned this Research*

Type of belief	Pragmatic beliefs and their roles throughout this research
Ontology: <i>The nature of reality</i>	Pragmatists believe that reality is situated within a certain social, historical, or political context, and that it cannot be separated from that context or the related interpretations of the knower. While an objective reality may exist, it cannot be separated out from subjective experience (Cherryholmes, 1992; Corbin & Strauss, 2008; Teddlie & Tashakkori, 2009) — hence the need for an interpretive stance with regard to analysis and interpretation in such research.
Epistemology: <i>The relationship between researcher and participant with regard to knowledge creation</i>	Pragmatists take an intersubjective approach, moving between the objective and subjective, believing them to lie on a continuum rather than being dichotomous, to find shared meaning (Morgan, 2007; Teddlie & Tashakkori, 2009). The relationship changes according to the research question or phase of the research (Teddlie & Tashakkori, 2009).
Axiology: <i>The role of the researcher's values</i>	The pragmatist researcher's values and interests determine what is to be studied (Teddlie & Tashakkori, 2009). They guide the research, shaping what is done and how it is done (Cherryholmes, 1992; Morgan, 2007). It is vital that researchers practise reflexivity throughout the research, to reflect on how their values, history, and experiences shape the research (Morgan, 2007, 2014).
Methodology: <i>The research process</i>	Methodological design is flexible and pluralist (Creswell & Poth, 2018; Creswell & Creswell, 2018). Pragmatic methodological decisions are emergent; the researcher responds and adapts to real-world constraints as the research unfolds (Patton, 2015b). Analysis is conducted abductively, with the researcher moving back and forth between deduction and induction (Morgan, 2007).

Research Design

This research utilised a sequential multi-method qualitative approach, combining several research and evaluation methods (Table 2). Multi-method research incorporates two or more different styles, methods, or orientations of research in one study (Brewer & Hunter, 2006; Hunter & Brewer, 2015). While mixed methods refer to the combination of quantitative and qualitative strategies, the multi-method approach allows any possible combination of methodologies (Creswell & Creswell, 2018; Teddlie & Tashakkori, 2009). As summarised in Table 2, this study incorporated different methods per research phase.

Qualitative Orientation

Braun and Clarke (2013) note that qualitative research can be critical or experiential. The approach used in the present study was experiential (or interpretive), which emphasises the research participants' own interpretations of their reality, with the researcher drawing on the meanings, views, perspectives, and experiences expressed by participants during the study.

The research objectives were achieved through a focus on my own and the research participants' experiences with the SAMEL Toolkit. For this reason, an interpretive lens was critical, as it aims to capture people's perspectives and prioritise these perspectives above those of the researcher (Braun & Clarke, 2013). However, the ultimate purpose of this research was not merely to understand but also to test a practical asset for use by the evaluation community. Following Goldkuhl (2012), an interpretive approach to *understanding* was combined with a pragmatic approach to *doing*.

While discouraged by methodological purists, combining qualitative paradigms benefits from a dialectical approach to inquiry that deepens one's understanding of a phenomenon (Greene, 2007; Maxwell, 2013). Evaluator Greene (2007) encourages one to

“invite multiple ways of knowing into the same study so that it may be deeply and generatively enriched” (p. 27). Similarly, qualitative research scholars Lincoln et al. (2018) note the increasing blurring of paradigmatic lines in qualitative research, highlighting the potential for “interweaving of viewpoints, for the incorporation of multiple perspectives, and for borrowing, or bricolage” (p. 228), a notably pragmatic goal.

Following the guidance of Goldkuhl (2012) in combining interpretivism and pragmatism, pragmatism was the base paradigm for this research, to serve the overarching aim of producing constructive knowledge for action. Interpretivism is considered instrumental in broadening the analytical view by incorporating research participants’ views, interpretations, and experiences.

The subsequent sections detail the methods applied in the two phases of the study, including sampling, data collection, and analysis.

Phase One: Toolkit Design and Expert Review

The objective of Phase One of this research was to design the SAMEL Toolkit. This entailed reviewing published peer-reviewed and grey literature primarily on MEL framework development and systems thinking but also literature on strategic programme management, organisational learning, and South Africa’s ECD sector. Drawing on these fields, I designed the Toolkit content and workshop framework, including guidelines for its administration. The development of the Toolkit is presented in Chapter Six. Thereafter, the Toolkit was disseminated amongst selected evaluation experts to sense-check and review its design.

Inviting experts to review draft documents is a suggested capacity-building technique in evaluations and MEL development (Peersman & Rogers, 2017). This was an important validation and quality assurance procedure to enhance the relevance and accuracy of the

Toolkit. Evaluators often use quality- and meta-evaluation standards to assess the design, implementation, reporting, and follow-up of an evaluation (Patton, 2012). While the Toolkit is not an evaluation, it is an experimental evaluation product.

To guide the quality review of the Toolkit, I developed a list of objectives that would indicate its quality (Step 2 in the Design Science Research process described above). These objectives would form the basis of the expert review (discussed below), and thus needed a meta-evaluative foundation. The objectives were formulated based on the African Evaluation Association (AfrEA) 2021 African Evaluation Principles (AfrEA, 2021). The review questions spoke to the AfrEA-based objectives, as well as the Made in Africa Evaluation Principles of Chilisa and Mertens (2021).

African Evaluation Principles and Made in Africa Evaluation

The African Evaluation Principles (AEPs) 2021 were developed between 2019 and 2021, under the auspices of the AfrEA, by volunteers from across the continent (AfrEA, 2021). They were developed to replace the African Evaluation Guidelines (AfrEA, 2007) which had been used for two decades to assess and improve the quality of evaluations in Africa. The AEPs are intended for use by all stakeholders involved in evaluation, to inform a high-quality, professional practice that is rooted in an African context (AfrEA, 2021). See Table 4 for a summary of the AEPs.

Table 4*Summary of the African Evaluation Principles 2021*

P: Powerful for Africans	T: Technically robust	E: Ethically sound	A: Africa-centric yet open	C: Connected with the world
P1. Conduct an appropriate, empowering process	T1. Be systematic and analytical T2. Be transparent and clear	E1. Be sensitive to stakeholders and relationships E2. Protect the rights of people E.3. Safeguard diversity and inclusion	A1. Engage with issues that matter in Africa A2. Consider framings and methods from Africa A3. Learn and adapt from the Global South, indigenous communities, and other contexts	C1. Acknowledge interdependence and interconnectedness C2. Foster the evaluation of sustainability in keeping with key international agreements and stewardship of nature C3. Strive to contribute to the urgent need for sustainable and transformative change
P2. Encourage reciprocity, including mutual accountability	T3. Be aware of dispositions T4. Ensure a feasible evaluation	E.4. Address inequalities and power asymmetries E.5. Be free from vested interests E.6. Consider trade-offs		
P3. Enable learning for useful insights	T5. Be efficient T6. Be culturally responsive			
P4. Value and strengthen domestic capacities				

Note. From “The African Evaluation Principles: 2021 Version”, by the African Evaluation Association, 2021, p. 5 (<https://afrea.org/AEP/new/The-African-Evaluation-Principles.pdf>).

The AEPs (AfrEA, 2021) reflect a growing evaluation movement in Africa that the AfrEA has championed: Made in African Evaluation (MAE). MAE is an Africa-centric approach that highlights African knowledge systems and practices that are culturally relevant and responsive to local contexts and communities' needs (Chilisa & Mertens, 2021; Dlakavu et al., 2022).

MAE represents a unified effort to rethink evaluations and pull away from the imperatives and standards of the Global North and donor influence (e.g., evaluations as objective, neutral, and accountability-driven), and the blind borrowing of Western frameworks, values, and standards to evaluate local programmes (Chilisa & Mertens, 2021; Masvaure & Motlanthe, 2022).

MAE is meant to complement global knowledge systems and perspectives while simultaneously uplifting African worldviews, theories, practices, and methodologies — all towards the goal of transformation and the betterment of society and nature (Chilisa & Mertens, 2021). As such, MAE authors and practitioners are working towards the decolonisation of evaluation in Africa and a methodology embedded in local context, one that empowers local communities (Masvaure & Motlanthe, 2022).

It was important to ground the SAMEL Toolkit (an African product), including its review, in this movement to ensure that it would contribute to a growing African practice and is aligned with the MAE movement's principles.

The criteria for the Toolkit were as follows:

1. The toolkit empowers the evaluator and the NPO team (AfrEA principle: *Powerful for Africans*);
2. The toolkit is technically robust (AfrEA principle: *Technically robust*);
3. The toolkit is ethically sound (AfrEA principle: *Ethically sound*); and
4. The toolkit has value for the evaluation field (AfrEA principles: *Africa-centric* and *Connected with the world*).

Phase One Sampling: Expert Reviewers

Purposive sampling, whereby participants are selected based on the researcher's judgement regarding their ability to provide in-depth knowledge and expertise (Braun & Clarke, 2013) was employed to select three experts to review the SAMEL Toolkit. One was an

experienced and knowledgeable international practitioner in the application of systems thinking to the field of evaluation, who was selected based on the person's substantial contributions to the systems thinking in evaluation literature. This expert was directly invited to take part in the review.

Purposive and snowball sampling were employed to identify an additional two South African experts to take part in the review. This approach involves drawing on the knowledge and expertise of key informants to identify eligible cases (Patton, 2015b). I consulted my personal networks, and those of my supervisors', to identify two experts experienced in MEL framework development for ECD programmes. The following selection criteria were used in selecting the South African MEL experts: a Master's degree or higher in programme evaluation, eight or more years of work experience in the field, and experience working in ECD. A profile of the experts, with their names replaced with pseudonyms, is provided in Table 5.

Table 5

Profile of Expert Reviewers

	Current location	Current profession	Education level	Years' experience in the field	Number of South African ECD projects	Familiarity with systems thinking in evaluation
Systems thinking expert (Sarah)	United States	Assistant professor	PhD	11 years in systems thinking and evaluation	N/A	Extensive: 20+ published works on the topic
South African MEL Expert 1 (Alletta)	South Africa	Monitoring and evaluation manager at an education consultancy	Master's degree	16 years in MEL	3	Self-reported as limited
South African MEL Expert 2 (Naledi)	South Africa	Senior evaluation consultant at an evaluation consultancy	Master's degree	18 years in MEL	6	Self-reported as moderate

Procedure and Data Collection: Expert Review

Once the experts had been identified, they were sent an invitation letter and consent form (Appendices A and B). Thereafter, they were sent the Toolkit, along with an introductory letter (Appendix C) and the review template (Appendix D).

The experts' review was prospective and formative in nature, as it necessitated looking at the design, plans, and procedures of the Toolkit (Harnar et al., 2020; Stufflebeam & Coryn, 2014). The review followed a semi-structured format, with the AfrEA guidelines and MAE principles used as guiding quality criteria (see Lynch et al., 2003).

The review template contained several review questions per Toolkit criterion, which probed the Toolkit in light of the AfrEA guidelines and MAE principles. For example, the Toolkit criterion *The Toolkit is technically robust* was aligned with AfrEA principle *Technically robust*, requiring that the evaluations be systematic, analytical, feasible, efficient, culturally

responsive, and transparent, while also taking predispositions into account (AfrEA, 2020). The review questions for this criterion/AfrEA principle were:

- Do you think that the Toolkit is sufficient to support evaluators and NPO teams to co-create a technically rigorous MEL framework? Why, or why not?;
- Are any of the steps missing important sources, evidence, concepts, or tools?; and
- How feasible are the toolkit steps in terms of typical NPO resources (human, financial, capacity and time)?

Experts were offered the choice of written feedback, using the template, or virtual oral feedback. All three chose written feedback, with two experts opting to provide additional in-text comments in the Toolkit documentation. To thank them for their time, the experts were given a small monetary gift.

Incorporating Expert Feedback

The post-review procedure involved reading all written feedback (including the in-text comments) twice and then consolidating it into a single document (see Figure 3).

Figure 3

Excerpt from Consolidated Expert Feedback

Goals	Review Questions	Sarah	Alletta	Naledi
<i>The toolkit empowers the evaluator and the NPO team.</i>	1. Please comment on this assertion: “the toolkit facilitates discovery, learning and agency for both the evaluator and the team”	Generally seems true – although it focuses more on the designers/leaders/commissioners than evaluators.	I concur that it does. Perhaps more so for the evaluator, who I understand would play a key role in facilitating discovery, agency and learning for the team. As a way to enhance discovery, learning and agency for the team, perhaps there could be certain sections (indicated with a different colour background for instance) which are primarily intended for the team. In some places the toolkit is relatively “dense” (for the non-academically minded) with terminology, concepts and a level of detail which is suitable for an evaluator (with intermediate knowledge/experience) but may be too advanced for some team members. Perhaps the text could distinguish between background, theory, tools and practice or highlight what is most relevant for team members vs for evaluators.	Agree with this. There is much emphasis on reflection and critical thinking about the design of programmes for both evaluator and the team. It also offers a good way to divide processes/tasks between programme team and evaluators, while still working collaboratively.
	2. To what extent do you think the toolkit can add to an	Some of what’s covered in the scoping and focusing will be new to evaluators, I suspect, and expand competencies although this	I think it can make a valuable addition/contribution, particularly in the	It helps with designing evaluations that are complex. The papers are also strong on

Thereafter, I summarised the feedback per goal, noting where experts agreed or disagreed, together with prevailing sentiments and suggestions. Finally, I re-read all the feedback and developed an actionable feedback matrix. I deemed feedback actionable if it provided a clear suggestion for a change or improvement that was feasible and practical, considering the study's time frame. All actionable feedback, per reviewer, was then grouped into five categories related to the Toolkit chapters: Scoping; Focusing; Monitoring & Learning; Evaluation; and a General category (miscellaneous comments on the design, format, or presentation). Finally, the comments in the matrix were ordered according to the content of each chapter, to ensure that comments pertaining to the same sections were grouped (see Figure 4).

Figure 4

Excerpt from Actionable Feedback Matrix

	Sarah	Alletta	Naledi
Focusing		<ul style="list-style-type: none"> - A real life ToC/ToA case example, guidance on how to deal with power dynamics during ToC workshops attended by a variety of stakeholders at different levels as well as advice/strategies on how to obtain agreement and approval to foster ownership. - Compare ToC checklist to an actual ToC (a good or bad one). - Survey stakeholder perspectives – could an example survey be provided? - A worked example of how to incorporate mechanisms and context into a ToC would be useful. - It would be helpful to provide guidance on how to deal with power dynamics during ToC workshops attended by a variety of stakeholders at different levels as well as advice/strategies on how to obtain agreement and approval to foster ownership. 	<ul style="list-style-type: none"> - The theory of change should depict the the multiple levels of the system that the programme will intervene with to bring about the impact envisaged. Often i find that the theoretical framework (like the SEM) needs to be visible within the TOC. - Discuss outcomes at different beneficiary levels – who are the ultimate beneficiaries? - Incorporate trigger questions into life space worksheets. Also in your tips also highlight that they should draw on the data in their situation analysis - Merge stakeholder analysis worksheet with situation analysis - Pre-workshop survey about ToC elements: shift focus from programme to system (in-text comment) - Combine mechanism and outcome workshop sections. - Present survey responses at the start of the workshop. - Greater attention to what is currently being done vs. what ought to be done (as identified by the workshop).

Deciding whether to incorporate the actionable feedback was a subjective process, with judgements guided by the following questions:

1. Is this suggested by more than one reviewer?;
2. Does the suggestion align with the Toolkit goals and purpose?;
3. Is the suggestion congruent with the core tenets of systems thinking?;
4. Does the suggestion markedly improve the quality of the Toolkit?; and
5. Is the suggestion pointing to a significant gap or problem?

Changes were made to the Toolkit design prior to moving to Phase Two of the research. These changes are described in detail in Chapter Six.

Phase Two: Testing the Toolkit

The objective of the second phase of the research was to test the SAMEL Toolkit with two ECD NPOs. Evaluators delivered the Toolkit content in five interactive workshops with the NPO management team (one full day per workshop).

Testing the SAMEL Toolkit was necessary to evaluate its performance capabilities, operational strategies and procedures, and overall design (Moore et al., 2011). Furthermore,

an important consideration for pragmatists is “what is beautiful and ugly and satisfying and dissatisfying and pleasurable and painful” to those for whom we wish to take action in a particular way (Cherryholmes, 1994, p. 17) – Toolkit users, in this case. Thus, identifying the Toolkit's pleasure- and pain points from the perspectives of the evaluator and participants (the users) was critical to understanding the Toolkit's feasibility.

The Toolkit was tested with two long-standing ECD NPOs based in Cape Town as test cases. I facilitated the first test case to ensure the Toolkit was implemented under ‘ideal’ circumstances (i.e., an evaluator who understands the content well, has read widely about systems thinking, is intimately familiar with the Toolkit, has a fair amount of time available, and has experience in the ECD sector).

The second test case was conducted by an experienced evaluator with no systems thinking experience or background, to ensure inclusion of ‘real life’ conditions: the evaluator is paid for his/her time, is less familiar with the Toolkit, is also busy with several other consulting projects, and has less experience in the ECD sector.

Testing the SAMEL Toolkit was a resource- and time-intensive exercise. Hence, the pragmatic decision was made to keep it to two initial iterations. The decision was taken that, should the two test cases’ outcomes be vastly different regarding the evaluators’ and the NPOs’ experiences, a third would be conducted with another NPO. This, however, proved unnecessary.

Phase Two Sampling

The Evaluator. Purposive sampling was used to recruit the evaluator who would conduct the test with the second case. This person needed to be based in Cape Town and available in the same period in which the NPO was available to conduct the test (which was the most difficult aspect to align). In addition, the evaluator needed to have at least a

Master's degree in programme evaluation and five years of experience in facilitating MEL development. This level of expertise was important because the SAMEL Toolkit was designed to be a complementary resource for evaluators already practising in the field. Given that it merges MEL with systems thinking, it can be argued to be advanced content for evaluators; it was not developed to teach new or emerging evaluators the basics of MEL.

Secondly, the evaluator needed to be well-versed in workshop facilitation, as facilitation techniques and basics were not included in the Toolkit. The focus of the test was the Toolkit's content, which necessitated expert facilitation.

I solicited an evaluation consultant from my professional network who I knew to be experienced in facilitating MEL workshops. I knew the person was working independently as an MEL consultant in Cape Town (which has a very small pool of evaluators) and was available during the testing timeline. The pseudonym 'Heather' (with confirmed 'she/her' as preferred pronouns) will describe her throughout the thesis.

The NPOs. Convenience- and snowball sampling were employed to identify eligible NPOs. The main criterion for participation was that the NPOs had to be ECD-focused. For practicality, the organisations had to be based in the Western Cape. Given the interactive nature of the Toolkit workshops and the time needed to facilitate the workshops, in-person field testing was necessary. Finally, the NPOs had to be interested and able to engage in the SAMEL workshops, and be willing to take part in the associated research activities (e.g., being observed, participating in focus group discussions, etc.). NPO teams also needed to be willing to revisit their programme strategies and commit the time and effort necessary to design a MEL framework. The judgement regarding their willingness was based on my initial interactions with the teams during recruitment. I did not consider this self-selection bias, as

the majority of NPOs who solicit MEL frameworks intentionally opt in and willingly commit the time and resources to the process, (although the extent of their commitment may vary).

The recruitment process was as follows. First, an advertisement was distributed amongst personal and professional ECD networks containing information about the research project and an invitation to participate (see Appendix E). I shared the advertisement on LinkedIn and via two ECD NPO membership networks: the National ECD Alliance and the BRIDGE Western Cape ECD Community of Practice. Second, the advertisement was shared amongst NPO members of the University of Cape Town's Knowledge Co-op (a network of NPOs who collaborate with university students on research projects). Interested parties were directed to an online form that assessed the extent to which they met the inclusion criteria and understood what participating would entail (see Appendix F). A total of seven NPOs responded to the advertisement.

Lastly, I employed snowball sampling by asking colleagues to recommend NPOs they thought might be interested in participating, whom I then contacted directly to share information about the research.

The final sample consisted of two NPOs, one selected from those who had responded to the advertisement, and one selected from those whom I had contacted. The directors of the two organisations were then sent an email formally inviting them to participate in the research, which also contained information on the study and a consent form (see Appendix G).

Test Cases

Characteristics of each organisation are summarised in the vignettes below (Figure 5 and Figure 6). The names of the organisations were replaced with the pseudonyms 'EduSpark' and 'EYEN'. At the time of the research, both organisations were going through a strategic

transformation period, hence their interest in this project. Both wanted to re-think their ECD offerings and develop new MEL frameworks. These were the primary reasons why they were selected to take part in the research — they were motivated to participate and excited about the process, and were willing to commit the time and personnel to attend all the workshops.

Both organisations had some experience in MEL, and had developed MEL frameworks in earlier years, but they did not have active MEL systems in place. With regard to MEL personnel, both organisations had undergone recent changes in staffing. One NPO had recently hired a new MEL manager, while the other had recently lost their MEL team.

I facilitated the testing of the SAMEL Toolkit with the first test case, EduSpark, while Heather facilitated the second, with EYEN.

Figure 5

EduSpark Profile



PILOT ONE:

EDUSPARK

INTRODUCTION TO THE ORGANISATION

EduSpark is an NPO that has been operating in the ECD sector for more than 20 years. The organisation is community-based, working in a single low-income area in Cape Town. It is relatively small, with just over 20 full-time and part-time employees, and is led by a CEO. The organisation partners with several other NPOs that offer ECD-adjacent services to the same community in which it works, and it has strong ties to the local government.

EduSpark offers three core programmes that span a child's educational lifespan, focusing on ECD, primary school education and secondary school education. The ECD programme is what formed the focus of the pilot. At the time of starting, the organisation's leadership were in the process of re-working the ECD programme's design.

The primary participants of EduSpark's ECD programme are teachers and principals in centre-based early learning programmes or preschools in EduSpark's community. The community is one of Cape Town's most vulnerable, and is known to be a particularly high-crime area with high rates of unemployment and poverty.

MEL AT EDUSPARK

At the time of the pilot, EduSpark's MEL capacity was being driven by each programme's management, with the assistance of two data capturers. Aside from the data capturers, no formal MEL personnel were in place. After the pilot, the data capturers left the organisation and full MEL responsibility was assigned to programme managers.

Prior to 2022, the organisation had a MEL team and MEL system firmly in place. In 2022, the MEL coordinator resigned and the team was dissolved. Since then, the organisation continued using the previous system to track reach and outputs in order to fundraise and report to donors and stakeholders. Data collection was primarily done offline using paper-based methods due to the safety concerns in the community, creating a backlog and significant delays in the organisation's MEL and reporting pipeline.

Due to the strategic re-working being done on the ECD programme at the time of the pilot, the organisation's management needed to create a new MEL framework and data collection tools.

The organisation's budget for MEL is approximately R340,000 per annum.

Figure 6

EYEN Profile



PILOT TWO:

EARLY YEARS EMPOWERMENT NETWORK (EYEN)

INTRODUCTION TO THE ORGANISATION

The Early Years Empowerment (EYEN) is an NPO that has been operating in the ECD sector for more than 40 years. The organisation has a sizeable footprint, working in three provinces across the country and has in excess of 140 employees and over 700 fieldworkers. It has an established reputation in the sector, and partners with a number of other NPOs, as well as government departments.

The organisation is led by a Chief Executive Officer and a Managing Director.

EYEN offers three core ECD-specific programmes, all operating in resource-poor contexts. One targets children's home environment via caregiver outreach, another supports playgroup programmes in communities that lack formalised preschool programmes, while the third supports more formal ECD centres in terms of training and resourcing so that they can register as partial care facilities with the DBE.

The primary participants of EYEN's programmes are primary caregivers that interact with children (i.e. parents and teachers), living in vulnerable and low-income areas (predominantly informal settlements and townships).

MEL AT EYEN

EYEN's MEL capacity is driven by one MEL manager, a data verification officer and two data capturers - all of whom are responsible for all MEL activities across all programmes. The MEL manager joined the organisation in 2022, and has no formal qualification in MEL. While some data collection was taking place at that time (MEL has been established in the organisation since 2015), no formalised MEL framework was being used.

When she joined the organisation, the MEL manager attempted to create logical frameworks for each programme using existing records and data collection tools, although gaps remained due to policy amendments and programme adaptations over time.

Prior to the pilot, MEL was primarily used in the organisation to track programme outputs, to track the organisation's reach and to report to donors and stakeholders. At the time of the pilot, the organisation was facing increasing pressure from donors to shift reporting to a more outcomes-based strategy, primarily focusing on the development of children. Additionally, the organisation's leadership wanted to engage in revised strategic work (including a new Theory of Change) to reflect ongoing adaptations that had been made to the programmes.

The organisation's budget for MEL is approximately R1,200,000 per year.

Phase Two Fieldwork and Data Collection

Each NPO participated in five in-person workshops, in which Heather and I explained the SAMEL Toolkit content and facilitated the activities. The workshop participants included programme managers and individuals involved in MEL activities. The aim was to hold one workshop per week, as the NPOs' calendars allowed; however, the timelines were interrupted by public holidays and participants' work commitments. EYEN had particular scheduling challenges, so their workshops were held every two to three weeks.

The test involving EduSpark took place off-site, at the University of Cape Town, due to safety concerns in the area in which EduSpark operates, while the second took place at EYEN's head office.

Data collection involved post-workshop focus group discussions, workshop observation, subjective reflections by the evaluators, and follow-up interviews with NPO directors (once the workshop series had been completed). All the workshops were recorded, and I personally transcribed^{vi} the recordings.

Focus groups. After each workshop, attendees participated in a focus group discussion on their experiences and thoughts regarding the day's content and activities. Focus groups were chosen as a pragmatic method to collect all participants' feedback in a short amount of time, in the same setting as the workshop, so that the experience would be fresh in their minds. The use of focus groups also allowed the feedback to emerge in a dynamic and interactive way, allowing participants to share multiple perspectives as a collective and in a

^{vi} Transcription was facilitated by the online artificial intelligence software Otter.ai. After the recordings had been transcribed automatically, I read the transcriptions while listening to the audio-recording, to ensure the accuracy of the transcripts. The software is password-protected.

non-intimidating environment (Braun & Clarke, 2013). For example, if a particular activity or concept was difficult to understand, it was hypothesised that participants would feel more comfortable sharing that with each other than in individual interviews.

Each focus group schedule (see Appendix H) was made up of different questions to keep each session less monotonous; however, they were structured similarly. Questions probed participants' biggest takeaways, ways in which their thinking had shifted, suggestions for improvement, as well as questions directed at the particular content covered in the workshop. Per the pragmatic paradigm, the questions were designed to be instrumental, collecting data on ways in which the Toolkit could be actioned and/or improved, as opposed to purely explanatory or narrative data (Cherryholmes, 1992; Goldkuhl, 2012).

Observations. Both sets of workshops were observed. A research assistant observed and recorded the workshops I facilitated, while I observed and recorded Heather's. In both cases, the observation was overt, but the observer did not participate in the workshops. The observations were substantive in nature, and included descriptions of the setting, the workshop activities and timeline, and behaviour of the facilitator and participants (Willig, 2013). Observations were deemed necessary to triangulate the data obtained directly from the participants and the facilitators, record the workshop events, and capture the participants' real-time reactions to and engagement with the workshop content and activities.

The observation schedule was semi-structured. Per workshop session, the observer recorded notes on the workshop activities, time taken, participant engagement, and the facilitator (see Appendix I). Under each of these sections, sensitising concepts were used to guide the observer and give some initial direction and organisation to the observations (Patton, 2015b), shown in Table 6.

Table 6*Sensitising Concepts for Workshop Observation*

Observation focus	Sensitising concepts
Participant engagement	<i>Interest, attention, questions, confusion, understanding, conversation</i>
Facilitator	<i>Pace, explanation, engagement, respect, inclusivity</i>

Evaluator accounts. During the course of the workshops, subjective information was collected on the experience of facilitating the SAMEL workshops. I recorded my reflections in an unstructured journal immediately after each of EduSpark’s workshops, while Heather provided feedback in the SAMEL Toolkit documents while working through the content, and via recorded informal conversations during the test (e.g., during lunch breaks and after wrapping up the day). Given the amount of work needed to deliver the workshops, additional self-observation and reflexivity work (e.g., journals or debriefing interviews) were deemed too burdensome for Heather. After Heather had completed the process, we engaged in a semi-structured final debrief interview (see Appendix J).

Follow-up interviews. Two months after completion of the workshops, I conducted follow-up interviews with the programme directors at both organisations. This was done to probe the strategic aftermath of the SAMEL workshops. As with the focus group schedules, the semi-structured interview schedule (see Appendix K) focused on pragmatic issues that would indicate action taken as a consequence of the workshops, i.e., how the teams made use of the MEL outputs that were developed, whether any changes were institutionalised, practical limitations of institutionalising the MEL framework, and practical intentions going forward.

Reflexivity: Situating the Researcher in the Research. Reflexivity is crucial in qualitative research, and involves critically reflecting on the process of knowledge production and the researcher's subjective role in that process (Braun & Clarke, 2013). Not only does this enhance the quality and rigour of a qualitative study, but it enriches the analysis process by embracing the researcher's subjectivity as a means to enrich the analysis and findings (Braun & Clarke, 2022). I embedded reflexive entries throughout all chapters of this thesis to clearly articulate my own positionality and influence over the research process. In the section below, I deliberate specifically on the fieldwork of Phase Two.

I, as the developer of the Toolkit, was also an evaluator/facilitator, observer, and analyst, and the effects of combining all of these roles should not go unsaid. The decision to facilitate one of the tests myself and observe the other was pragmatic (to save costs), ideal (in terms of knowledge and expertise), and aided reflexivity — I wanted to embed myself in the process, from beginning to end, as much as possible.

In the first test, I wore two hats: developer and facilitator. This meant I had a higher level of expertise than a typical evaluator would when using the SAMEL Toolkit, and it also meant that I was personally and professionally invested in the outcomes of the process. I wanted it to go well. This would have influenced the participants' experience of the workshops and the overall quality, as I put a lot of time and effort into planning the workshops and developing the MEL products, perhaps more than an average evaluator would.

In the second test, I again wore two hats, this time of developer and observer. Again, I wanted it to go well, but I was also acutely attentive, vigilant, and critical. Further, I was comparing Heather's test to my own. Due to these qualities, my observation notes were long and detailed, far more so than the notes taken by my observer in the first test. Finally,

Heather delivered the workshops in the presence of the developer of the Toolkit, placing her in an intimidating position. While she mentioned that this made her nervous at first, she appeared to relax quite quickly, and she then used my presence in the room to her advantage, asking questions or seeking advice. Thus, while the second test was intended to be more related to the 'real world' than the first, my presence as observer certainly influenced the process, even though I tried to remain as uninvolved as possible.

The final dual role worth noting here is that of developer/focus group moderator. The focus group participants in the second test were less critical of the Toolkit and the overall process, compared to the first test group. This is likely because they felt uncomfortable providing criticism to the Toolkit developer (they knew I had developed it). However, constructive criticism was low in the first test as well. I believe that this is possibly due to gratitude. Both groups continuously expressed gratitude for the process throughout the workshops, and may not have wanted to 'bite the hand that fed them'.

Phase Two Data Analysis

Two approaches were taken to analyse and report on the test data. First, a combined case description was developed to provide a narrative description of the test proceedings, as well as the evaluators' and NPO teams' engagement with the Toolkit content. Case descriptions are appropriate when the research goal is descriptive in nature, and, thus, an inductive or deductive approach is unsuitable (Yin, 2018). As the SAMEL Toolkit had never been implemented before, it was important to describe its first-time delivery without applying an interpretive lens. Case descriptions are typically organised around topics (Yin, 2018). The five SAMEL workshops were used as descriptive topics (i.e., Scoping, Focusing – Theory of Change, Focusing – Theory of Action, Monitoring & Learning, and Evaluation). Each NPO case was used as a sub-topic. After an initial period of data immersion, the focus group

transcripts, observation notes, journals, transcripts of Heather's 'check-ins', and interview transcripts were sorted into these topics. Thereafter, a description of the test was developed, presented in Chapter Eight.

Following the test description, I conducted reflexive thematic analysis on the same data corpus to generate inductive themes in the data that linked with the NPOs' experience of the SAMEL workshops. This approach allowed me to generate patterns in the data by applying a more interpretive lens to understand how the NPOs experienced the Toolkit.

Reflexive thematic analysis is a systematic approach to identifying, analysing, and reporting themes or patterns within a dataset (Braun & Clarke, 2006, 2013, 2022). It is known to be a flexible approach that can be used both deductively or inductively (Braun & Clarke, 2006, 2013, 2022). This approach was chosen due to its flexibility. An initial inductive analysis meant that the themes that I developed would be grounded in the experiences of the test participants; thereafter, I could take a deductive approach to interpretation, in which I could understand the themes using the literature. Furthermore, an analysis based in patterned meaning allowed me to identify commonalities in experiences across the two test cases.

Given the flexible nature of reflexive thematic analysis, it is important to locate the analysis approach within the broader research paradigm and epistemological and ontological orientation of the study (Terry et al., 2017). I employed thematic analysis under the pragmatist banner of contextualism, in which reality is situated within the historical context of the researcher and participants (Braun & Clarke, 2013). As such, the findings vary according to the context in which the data were collected and analysed (Madill et al., 2000).

The aim is to understand the truth (a realist position), but both contextualism and pragmatism hold that there is no single truth that is universally applicable. Rather, the researcher attempts to represent participants' perspectives by drawing on their descriptions

and interpretations (Madill et al., 2000). Thus, the aim is understanding what is true in the particular context under study (Tebes, 2005), and an interpretive approach (discussed earlier) is appropriate. Contextualism does not require one to justify or interrogate participants' views, but to present knowledge claims in light of their context (Tebes, 2005).

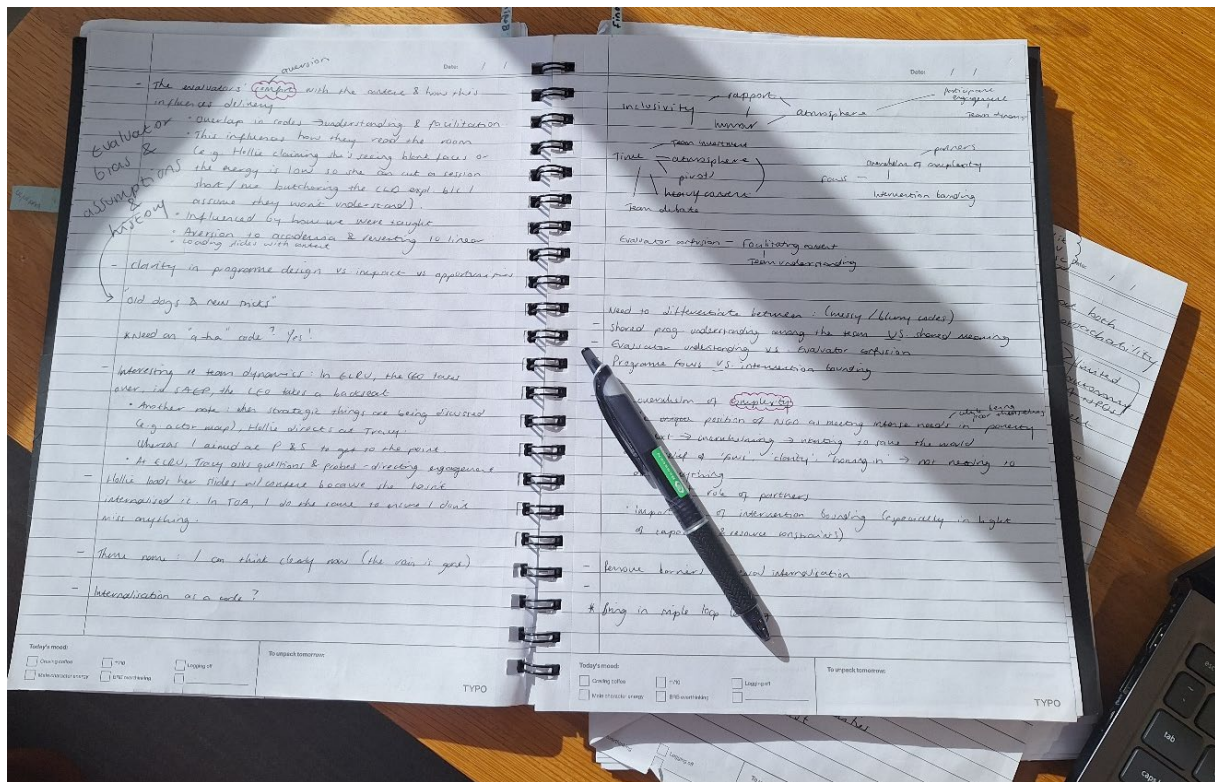
Taken to the extreme, this approach could result in nonspecific findings that cannot be applied outside of the current context (Tebes, 2005). However, as the present study was not based in hypothesis or causal claims, the contextualist view does not limit its validity or utility. Rather, it allows us to understand the relationship between the SAMEL Toolkit and its implementation setting, which is key in learning about feasibility issues in real-world settings.

Thematic Analysis Steps. I followed the six-step approach to reflexive thematic analysis recommended by Braun and Clarke (2006, 2013, 2022): (1) familiarising oneself with the dataset; (2) generating initial codes; (3) collating codes into potential themes; (4) reviewing themes using thematic maps; (5) defining and naming themes; and (6) reporting the findings. While these are presented as distinct steps, Braun and Clarke (2006, 2013, 2022) note that the steps are iterative and non-linear. With the exception of data immersion, I involved my supervisors in all phases of the analysis, to ensure that I was not interpreting fragments of data in isolation. Given my involvement in the data collection process, it was important to include outsider perspectives in the analysis.

In Step 1, I immersed myself in the data by reading all the transcripts, observations, and journal entries twice. I was already quite familiar with the data due to my involvement in the collection process, as well as the transcription. As illustrated in Figure 7, this was an 'active' reading of the data, which included taking notes, generating analytic ideas and insights, and making observations regarding information contained in the data (Braun & Clarke, 2006, 2013, 2022).

Figure 7

Illustration of Active Reading of the Data

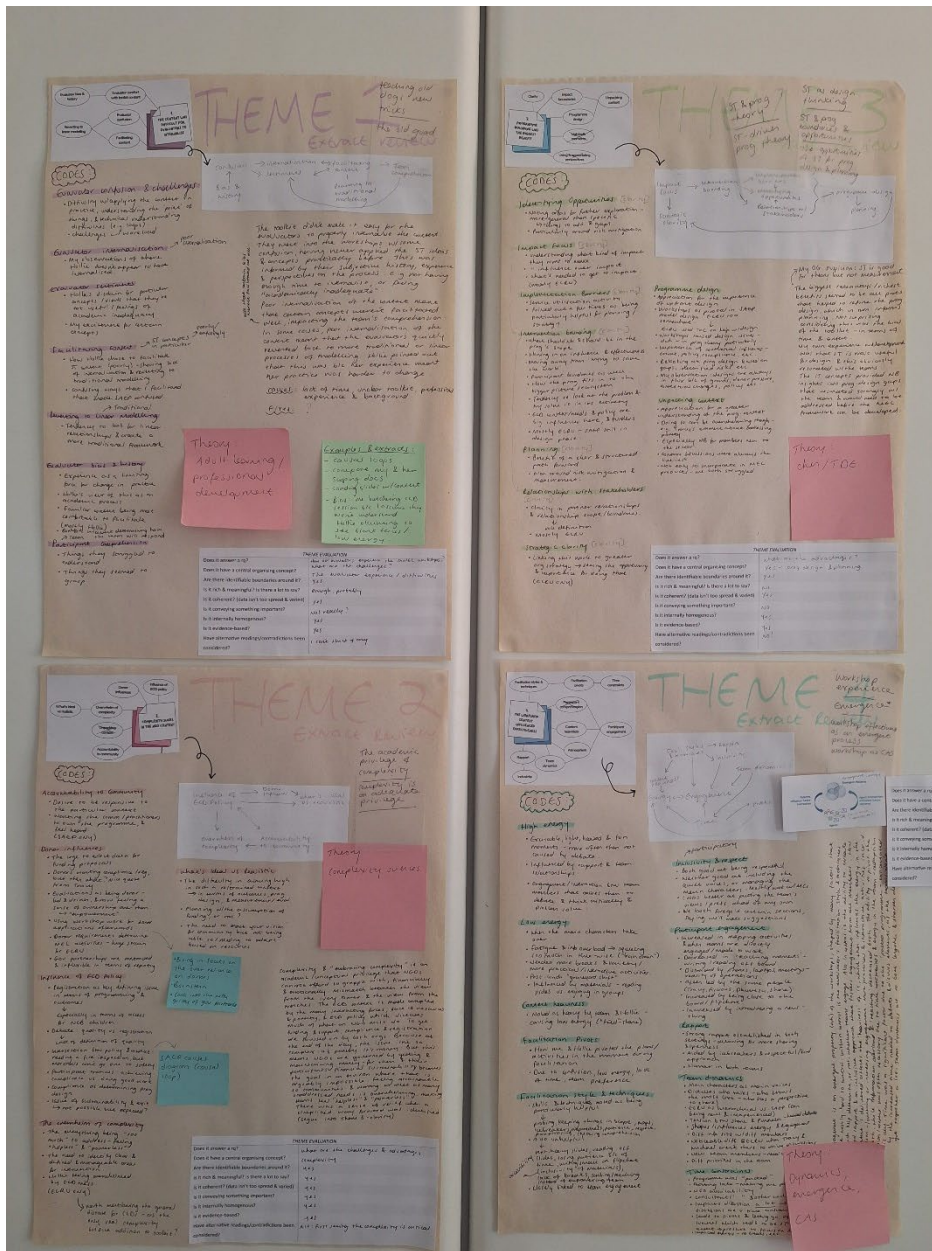


In Step 2, I transferred all textual data items into NVivo 14, a computer-assisted qualitative data analysis software (CAQDAS) package. Using the software, I began generating semantic codes using the research questions as guides — reflecting what participants explicitly stated and how they made meaning of their experience (Braun & Clarke, 2006, 2013, 2022). My coding initially produced 64 codes, but many overlapped. The codes were refined and reduced during subsequent steps in the analysis (see Appendix L for code descriptions).

In Step 3, I developed candidate themes by actively looking for patterns amongst the initial codes (Braun & Clarke, 2006, 2013, 2022). I looked for similarities and relationships amongst the codes to identify a central organising concept, using thematic maps (Braun &

Figure 9

Brainstorming Candidate Themes



In Step 4, I reviewed the candidate themes and thematic maps to establish how well they reflected the whole dataset and the coded data, and whether they answered the research questions (Braun & Clarke, 2006, 2013, 2022). First, I looked at the data extracts per theme to establish how well the theme and its central organising concept captured the

meaning in each coded extract. I then reflected on the dataset as a whole to determine whether the themes fit the dataset. Thereafter, I considered the research questions to assess whether the themes provided answers. Finally, I engaged in a quality-checking exercise per theme, where I asked:

- Is there external heterogeneity? (Patton, 2015b). Is this theme distinctive from other themes in terms of its central organising concept? Does it have strong or blurry boundaries? How many codes belong to this theme and to others?
- Is there internal homogeneity? (Patton, 2015b). Do the associated data extracts cohere meaningfully?
- Is the theme related to the other themes? Does it contribute to a cohesive narrative?

This process was highly iterative, with the themes changing several times as I continually referred to the data corpus and engaged in sense-checking. I involved my supervisors in this final phase, to obtain an outsider perspective.

In Step 5, I finalised the themes by naming them and then formulating succinct descriptions of each by drawing on the associated data extracts. To develop the descriptions, I asked the following:

- What is interesting about this theme and set of extracts? Why?
- How does this theme answer the research questions?
- What part of the greater story does this theme tell?

See Appendix M for descriptions of the themes.

Finally, in Step 6, I produced the analysis write-up, which is presented in Chapter Nine.

Ethical Considerations

Ethical approval for this research was sought and obtained from the Ethics in Research Committee of the Faculty of Commerce (see Appendix N). A data management plan was developed prior to embarking on the research (see Appendix O).

The Code of Ethics of the British Psychological Society (2018) is based on four guiding principles that should be considered and applied throughout the research process: respect, competence, responsibility, and integrity.

Participants (including the NPO participants, expert reviewers, and the evaluator) voluntarily participated in this research, and were informed that they could opt out with no consequences. All test participants, including Heather, signed consent forms (see Appendix P) before the research commenced, which explained: the purpose of the research; the role they would play, should they choose to participate; what participation would entail; that their participation would be recorded and transcribed; that participation was voluntary; that there would be no cost implications (for NPOs); and that their identities would be protected by using pseudonyms in the reporting (in adherence to the Protection of Personal Information Act 4 of 2013). Verbal consent was sought during the workshops to include de-identified photographs in this thesis. Heather, my research assistant, and the expert reviewers signed non-disclosure agreements developed by the university's Research Contracts and Innovation Department.

All data pertaining to this research were stored in a password-protected cloud on password-protected devices.

Despite this research posing no harm to participants, there was a financial incentive for participation. Expert reviewers were paid an honorarium as a token of appreciation for their time, while Heather was remunerated for services as workshop facilitator. The NPOs

were given the workshop process and MEL products for free. While this could be considered coercive, Heather and the expert reviewers indicated a keen interest to participate before knowing about the remuneration. All expressed an interest in the purpose and outcomes of the research, and were willing to contribute their expertise. Regarding the NPOs, the research relationship can be deemed mutually beneficial. While the teams' participation was time-consuming, they received what is generally an expensive strategic process for free.

Finally, in terms of my own professional competence, responsibility, and integrity, I intentionally and continuously engaged in reflexive practice throughout the research, considering my own biases, skills and competencies, assumptions, and personal history. This was particularly important during the tests, as I observed the Toolkit come to life while grappling with my various roles in the research (noted earlier). Throughout this thesis, I surface my reflexivity, particularly my relationship with and lived experience of the topic, the chosen research methods and associated pragmatic beliefs, and my roles in the data collection process.

In the following chapter, I describe the development of the SAMEL Toolkit.

Chapter Six: The Development of the SAMEL Toolkit

The Systems Approach to MEL (SAMEL) Toolkit is a workshop-based systems approach to developing an MEL framework for an NPO. The version of the Toolkit developed for this research was focused on NPO evaluators or MEL practitioners^{vii} working in NPO ECDs. The Toolkit introduces evaluators to existing systems thinking tools and concepts, and provides guidance on practically applying the content to the development of MEL frameworks. Practical application of the Toolkit takes place via a series of five facilitated workshops with NPO staff, in which the evaluator uses the Toolkit contents to collaboratively develop a MEL framework informed by systems thinking.

In this chapter, I introduce the Toolkit, and then present the conceptual background that guided its development. Thereafter, I describe my development approach using bricolage. The content of the Toolkit is described in detail in the next chapter.

Description of the SAMEL Toolkit

The SAMEL Toolkit is a guiding document consisting of five chapters: (1) Introduction; (2) Scoping; (3) Focusing; (4) Monitoring & Learning; and (5) Evaluation. Toolkit chapters two to five describe the components of the MEL framework that would be developed in the workshops (see Table 7), including an introduction to the systems thinking concepts and tools to be used to develop these components. Each chapter also contains links to worksheets and activities to practise the tools described in the chapter, as well as step-by-step guidelines for facilitating each workshop.

^{vii} I use the term 'evaluators' hereafter to refer to the SAMEL Toolkit users, both those who participated in this research and prospective users.

Table 7

SAMEL Outline

Chapter (Workshop)	MEL components (Workshop outputs)
Scoping (Workshop 1)	Situation analysis Programme scope (problem statement, vision, mission, values)
Focusing (Workshops 2 and 3)	Theory of Change Theory of Action
Monitoring & Learning (Workshop 4)	Monitoring framework Learning strategy
Evaluation (Workshop 5)	Ideal evaluation plan

Toolkit Audiences

The SAMEL Toolkit has two audiences. The primary audience is the evaluators who engage with the Toolkit directly, and for whom the Toolkit was developed. The secondary audience is the NPO team members who participate in the SAMEL workshops delivered by the evaluator. I knew the experience and knowledge of MEL amongst the two groups would differ substantially, and that these differences had to be incorporated into the Toolkit's design.

I developed the Toolkit content to appeal to a variety of evaluators with varying levels of training and experience. However, I operated under the assumption that users would have a fundamental understanding of MEL and some practical experience in developing MEL frameworks. The Toolkit is not intended to introduce MEL, but rather to act as a complementary resource to traditional texts, thus building systems thinking competencies on an existing MEL knowledge base.

Regarding the NPO teams, the Toolkit embeds the systems thinking content in the SAMEL workshop activities and processes rather than delivering instruction on systems thinking principles. The rationale here was to keep teams focused on MEL and use systems thinking as a lens through which MEL capacity is developed.

While the SAMEL Toolkit was initially developed as a test product tailored for African evaluators working in the ECD sector, I believe its content and tools are applicable to evaluators working in any sector or region. The Toolkit incorporates systems thinking concepts and practical applications that address universal challenges and priorities in MEL, such as the need for stakeholder engagement, context-sensitive programme design, and iterative learning processes. Moreover, while ECD is used for illustration purposes, the majority of the Toolkit content is sector-neutral.

Toolkit Design

The SAMEL Toolkit is an instructional learning resource for evaluators. To maximise its potential effectiveness, I designed the Toolkit using principles from the field of instructional design. Instructional design focuses on the systematic development of educational and training content and programmes to ensure efficient, effective, and appealing learning experiences (Brown & Green, 2016; Seel et al., 2017). The field synthesises knowledge from various disciplines, including cognitive psychology, pedagogy, and communication studies to design, implement, and evaluate instructional materials and systems (Brown & Green, 2016; Seel et al., 2017).

Drawing on pragmatist philosophy, Parrish (2009) emphasises aesthetics—the sensual characteristics of the designed object or experience—as a way to create particularly immersive and meaningful learning experiences (see also Uhrmacher, 2009). Parrish (2009) provides five aesthetic principles for instructional design; these are listed in Table 8, together with how they were applied in designing the Toolkit.

Table 8*Applying Instructional Design Principles to the Toolkit*

Parrish's aesthetic principles of instructional design		My application
1: Learning experiences have a beginning, middle, and ending (i.e., a plot).	Learning experiences should be structured with a narrative-like sequence, akin to the plot of a story.	The SAMEL Toolkit's introductory chapter poses the relevant 'problem' of linear approaches to MEL to create a sense of tension and an interest in systems thinking as a solution. The Toolkit structure follows a coherent, scaffolded 'story' that starts with the programme situation and ends in evaluation. Each chapter builds on the content and narrative of the preceding one.
2: Learners are the protagonists in their own learning experiences.	The learning experience should be personalised and relevant to the learners, allowing them to actively engage and see themselves as central to the learning process.	The SAMEL Toolkit is a self-paced learning resource; evaluators can determine how and when they engage with the content. Worksheets are provided as a way for evaluators to practise what they learn and reflect on their understanding of concepts.
3: Learning activity, not subject matter, determines the theme of instruction.	The activities through which learners engage with content/subject matter are more important than the content itself, necessitating engaging activities that are connected to experience.	The toolkit uses real-world examples to explain concepts, and worksheets are designed to help users practically apply what they learn.
4: Context contributes to immersion in the instructional situation.	The physical and psychological context in which learning occurs significantly affects the learner's engagement and immersion.	While the physical learning environment is not addressed as part of the Toolkit, I designed the Toolkit to be as visually appealing and engaging as possible. Colour-coding, graphics, illustrations, and clean design principles support ongoing engagement with the material.
5: Instructors and instructional designers are authors, supporting characters, and model protagonists.	Instructors and designers should be models of engaged, curious learners. They should remain reflexive and transparent in their own learnings, and designers should use empathy in their design.	I relied heavily on my multifaceted role — researcher, systems thinking novice, evaluator, and Toolkit designer — to develop the SAMEL content. I deliberately introduced and presented the content in a way that would prevent the frustration, confusion, and overwhelm I had often felt when first engaging with systems thinking literature.

SAMEL Conceptual Background (1): The Systems Concepts

The development of the SAMEL Toolkit content was guided by the systems concepts approach to systemic evaluation (described in Chapter Three). I opted to blend the DSRP theory (Cabrera et al., 2015; Cabrera & Colosi, 2008) and SETIG principles (SETIG, 2018) to arrive at the following concepts: *systems*, *boundaries*, *relationships*, and *perspectives*, summarised below.

1. There are clear distinctions both between and amongst systems and sub-systems.

Boundaries define what is included and, equally important, what is excluded.

2. All **systems** consist of identifiable parts (things and ideas) that are organised in a particular way to form an identifiable whole and to achieve a particular purpose or function.

3. There are meaningful **relationships** between and amongst systems, sub-systems, and system parts, which determine how they function and interact with each other. These relationships determine emergent dynamics between system parts.

4. Different people have different **perspectives** that inform how the meanings of system parts and wholes are constructed.

As stated in Chapter Three, I agree with Patton's (2023) pragmatic positioning of the systems concepts, that they should be an approachable and easily understandable application of systems thinking. Furthermore, I believe that the systems concepts, while a product of Western epistemology, align well with MAE (discussed in Chapter Five).

Systems Thinking and MAE

As discussed previously, the foundational ethos of systems thinking – recognising interconnectedness and interdependence – has long been intrinsic to indigenous knowledge systems, including African philosophies like Ubuntu and Ukama (Chilisa & Mertens, 2021;

Njenga & Massyn, 2024). To align with and contribute to the MAE movement, I noted the importance of grounding the SAMEL Toolkit in its African context, however, this posed a significant challenge in terms of the availability of practical tools and conceptual frameworks. While African evaluation paradigms and approaches are well-established theoretically, their practical application often relies on integrating Western methodologies due to a lack of practical guidelines (Fish, 2022).

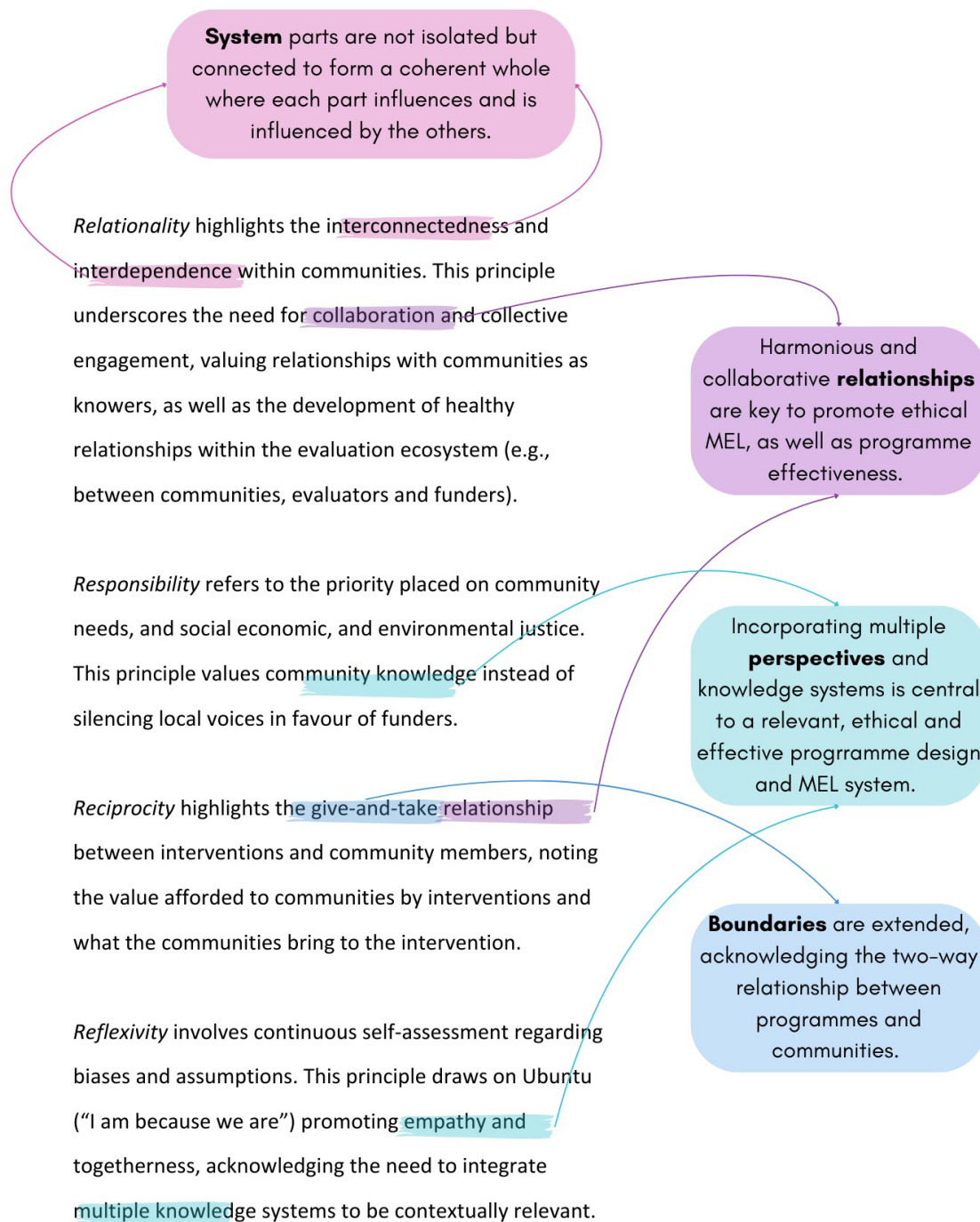
Thus, while the SAMEL Toolkit aligns with the principles of MAE and seeks to incorporate decolonial perspectives, it also draws heavily on systems thinking tools and concepts that originated in Western academic traditions. Figure 10 overleaf illustrates the connections between four MAE principles (Chilisa & Mertens, 2021) and the systems thinking concepts.

This tension presents both an opportunity and a challenge. On the one hand, it offers a chance to integrate the strengths of established and tested Western systems thinking tools with the relational and holistic epistemologies of African knowledge systems – which is encouraged in MAE (Chilisa & Mertens, 2021). On the other hand, it risks perpetuating the dominance of Western paradigms unless these tools are critically adapted to reflect and respect the cultural and contextual realities of African contexts.

I believe that this tension is a necessary part of the evolution of systems thinking in evaluation. By using Western tools as a starting point and contextualising them through African philosophies and participatory methods, the SAMEL Toolkit contributes to the growing body of work that seeks to integrate Western and African knowledge systems. For example, in Chapter Seven, I describe how systems thinking tools complement relational African approaches to evaluation to create a more inclusive and contextually responsive evaluation framework.

Figure 10

Alignment Between MAE and Systems Concepts



In applying the systems concepts, I also felt it was important to address some concerns regarding the concepts, particularly their vulnerability to being misused due to different interpretations by different users (Jackson, 2023a; SETIG, 2018). Additionally, critics have argued that various system methodologies and concepts are understood and applied differently (Jackson, 2023a; Miller, 2016). For example, the concept of *boundaries* has real-world or physical implications in systems dynamics, whereas in Critical Systems Heuristics (Ulrich, 1983), boundaries are applied in discussions of social values, ethics, and justice.

Addressing Critique of Systems Concepts

In a critique of Patton's (2011, 2015a) use of the concepts in developmental evaluation, Miller (2016) argues that sensitising concepts (as Patton conceptualises the system concepts) are not enough: "The practitioner must learn what systems concepts mean and how to carry out the methods that go with them, so these can be mindfully adapted to the evaluation context" (p. 268).

I agree with Miller's (2016) view, which is why a major objective in designing the SAMEL Toolkit was to incorporate the systems concepts in meaningful and applicable ways. This was done by providing specific instructions for using the concepts in the context of MEL facilitation.

Firstly, the systems concepts were uniquely applied to each part of the SAMEL process. As such, the meanings of the concepts were adapted to the particular MEL component being workshopped. For example, *relationships* could refer to the NPO's partnerships with other organisations (a component of the Action Model in Chen's [2015] schema), or causal relationships espoused in a programme theory (a component of Chen's [2015] Change Model). The Toolkit's application of the concepts thus changed according to the workshop and associated MEL component being developed.

Secondly, I drew on systems thinking methodologies to identify specific tools that would help evaluators apply each of the concepts in each workshop context (this is described in full in the following chapter). Actor mapping was the tool used to identify the NPOs' external relationships, while feedback loops and system archetypes were used as tools to identify non-linear relationships in the programme theory. In this way, I attempted to provide evaluators with a clear understanding of how the systems concepts could be adapted to different MEL components, while also providing existing systems thinking tools to aid accurate application of the concepts.

This approach was my own attempt to find Jackson's (2023b) "golden mean" (p. 637) of systemic evaluation. As discussed in Chapter Three, Jackson (2023b) argues for the need to find a balance between systems concepts (minimum specifications) and the more in-depth single systems methodology approach (maximum specifications). The Toolkit does this by defining the systems concepts per MEL component and then matching those definitions with particular systems thinking tools from single methodologies. The rationale for doing so aligns with the views of Patton (2023) and Williams (2019): it is not practically feasible to expect newcomers to systems thinking to learn all of the methodologies that might be suitable for different evaluation situations and evaluands. It is critical to explore ways to support evaluators in discovering the benefits of systems thinking while removing the theoretical barriers that may put them off this discovery from the outset.

As stated in the previous chapter, this research was based in pragmatism. It is my opinion that methodologies that represent maximum specifications, like Critical Systems Practice (Jackson, 2023b), are unlikely to be taken up by evaluators in the time available to transform evaluation to suit an increasingly complex world. To transform, evaluation needs

more pragmatists, like Patton, who understand the real-world constraints of evaluation. Miller (2016, p. 268), in a critique of Patton's developmental evaluation, noted the following:

Patton is among our most appealing theorists because his theoretical work reflects the careful integration of his experiences in practice gained over years of trial. He attends closely to what his clients find helpful. He uses his practice as the place to discover and test emerging ideas. He pays close attention to practitioners' experiences applying his ideas. He aspires to get it right and say it clearly. He writes with the larger goal of helping practitioners and their clients get the best possible evaluation for their particular needs.

It is in Patton's footprint that the SAMEL Toolkit was developed and tested.

SAMEL Conceptual Background (2): Theory-based Evaluation

In addition to the systems concepts, the Toolkit was based on a programme theory approach to MEL. Chen (2015) and other evaluation scholars (e.g., Donaldson, 2007; Funnel & Rogers, 2011; Rossi et al., 2019) argue for the utility of theory-based evaluation (TBE) in unpacking causal relationships and answering *how* and *why* questions about programme success. This approach requires that evaluators examine how programme stakeholders' assumptions about the design, implementation and outcomes of the programme (known as 'programme theory') operate in the real world, and whether they accurately represent how programmes succeed in their goals (Chen, 2015).

Programme theory lays the foundation for monitoring and evaluating participant outcomes and programme implementation (Rossi et al., 2019). Measuring programme participants' outcomes indicates how effectively a programme has achieved its intended benefits (Rossi et al., 2019). Pairing outcome evaluation with an investigation of the programme's implementation further helps to explain *why* a programme succeeds or fails,

and how outcome data should be interpreted (Durlak & DuPre, 2008; Dusenbury, 2003). Additionally, understanding the barriers to implementation could inform programme adaptation, where necessary, and improve the service being delivered (Patton, 2012; Rossi et al., 2019). Thus, we can use this information to understand how and why programmes are, or are not, effective.

Critics of TBE argue that developing programme theory can be resource and time-intensive, also requiring significant expertise to do properly but with limited available guidance on a theory's validity or credibility (Coryn, 2005; Coryn et al., 2011). In addition, if evaluators consider a programme to be lacking a robust and comprehensive programme theory, they may attempt to develop a theory (which may turn out to be unsound) within time- and resource constraints, and thereby "usurp the program staff's responsibility for program design" (Stufflebeam, 2001, p. 39). Finally, critics have argued that evaluators may overemphasise the evaluation of the theory itself instead of the programme (Coryn, 2008).

TBE was used as the second conceptual framework for the SAMEL Toolkit for several reasons, several of which provide a counterargument to the points above. (1) Most NPOs in South Africa are expected by funders and stakeholders to have some form of programme theory (Wildschut, 2014). It is, therefore, common that the time and resources are made available for the exercise. (2) In my consulting experience, I have found programme theory to be a useful and comprehensible MEL foundation for my NPO clients if developed in a participator manner (as proposed in the Toolkit). Involving programme staff in the process supports their active engagement with programme design. (3) I believe that systems thinking is well suited to programme theory development, and the two work well together as a strategic foundation for MEL (illustrated in detail in Chapter Seven). (4) Several research studies have found ToC to be particularly helpful in complex programme settings when

combined with systems thinking tools (e.g., Chapman et al., 2023; Douthwaite et al., 2020; Douthwaite & Hoffecker, 2017; Wilkinson et al., 2021).

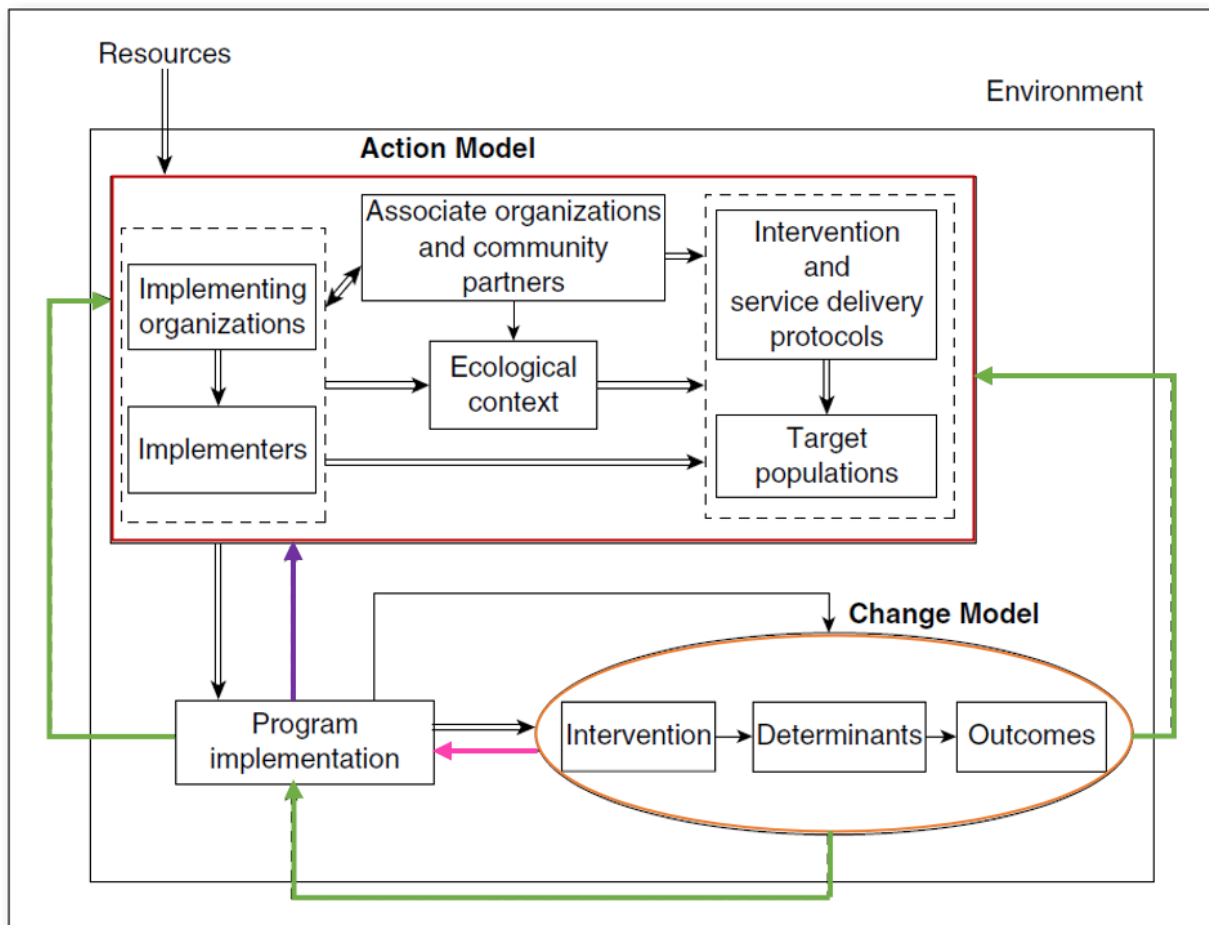
I drew on Chen's (2015) Action Model/Change Model Schema as the guiding TBE framework for the SAMEL Toolkit.

Chen's Action Model/Change Model Schema

Chen (2015) views a programme as an open system in which inputs are transformed into outputs (programme results), and posits that the external environment plays an important role in influencing this transformation process. Chen (2015) developed a conceptual framework to guide evaluation that incorporates this view: the Action Model/Change Model Schema, shown in Figure 11. The Action Model captures stakeholders' prescriptive assumptions — what actions they believe they need to take to produce the intended programme results. These assumptions determine how exactly a programme should be implemented to be effective. The Change Model captures stakeholders' descriptive assumptions — what causal processes they expect to take place to achieve the desired participant outcomes and intended impact. To achieve this change, Chen (2015) asserts, the Action Model must be implemented effectively. The Action Model, Change Model, and feedback are illustrated in Figure 11 and discussed below.

Figure 11

The Action Model/Change Model Schema (Comprehensive Form)



Note. Adapted from *Practical Program Evaluation: Theory-driven Evaluation and the Integrated Evaluation Perspective* (p. 81), by H. T. Chen, 2015, Sage Publications. Copyright 2015 by Sage Publications.

The Change Model. Chen’s (2015) Change Model (circled in orange in Figure 11) consists of the actual programme or intervention, determinants, and outcomes. Outcomes are the intended change that the programme is meant to achieve in its target population. These outcomes are brought about by determinants, which are mediating or intervening variables that the programme stakeholders believe contribute to the underlying problem. Addressing these variables lays the foundation for achieving the desired outcomes. Thus, the

Change Model depicts how the programme and its activities address determinants (or causes of the problem), which, in turn, lead to the intended outcomes.

The Change Model is a part of what Chen (2015) refers to as the 'programme scope', which is a statement of the programme's focus and goals. The programme scope incorporates: the problem being addressed; the target population that requires the intervention; and the Change Model, which illustrates how the programme's intervention is expected to achieve its goals. Chen's (2015) programme scope describes what is more commonly referred to as a programme's ToC (Funnel & Rogers, 2011). In their book *Purposeful program theory*, Funnel and Rogers (2011) refer to the constituents of ToC as: a situation analysis (What is the problem, its causes, and its consequences?); defining the boundaries of the programme; and an outcomes chain, which visually illustrates the causal relationships between the intervention and its short-term and longer-term outcomes, akin to Chen's (2015) Change Model.

The Action Model. The Action Model (Chen, 2015) (boxed in red in Figure 11) represents the broad implementation plan for delivering the programme. It consists of the organisation implementing the programme and its capacity to do so; the programme implementers and their competence in and commitment to carrying out the work; associate organisations and community partners that offer complimentary services; the ecological context that supports the programme at a micro and macro level; the intervention and service delivery protocols that guide programme delivery; and the processes and strategies for engaging the target population.

Feedback. The coloured arrows in Figure 11 depict evaluation feedback and how it informs the Change Model and the Action Model (Chen, 2015). Feedback can be internal, relating to the monitoring data that demonstrate to programme personnel how effective the

current implementation is and what changes need to be made to the Action Model to improve implementation (purple arrow). Additionally, information from the Change Model (on participants' outcomes) can be used to inform the implementation of the programme (pink arrow). External evaluation feedback extends beyond the boundaries of the programme, to include interested stakeholders and decision-makers such as funders and the general public (green arrows). This feedback illustrates the merits of the programme, what improvements need to be made, and what future direction the programme should take.

Chen's Schema and the SAMEL Toolkit

Chen's (2015) Action/Change Model Schema is a useful and intuitive guide in MEL development. Specifically, Chen (2015) clearly articulates the ongoing and dynamic relationship between programme outcomes and implementation, as well as the influential role of the programme environment (and feedback). However, while Chen (2015) acknowledges that programmes are open systems, the guidance for the development of Action/Change Models does not include ideas or concepts from systems thinking. This is a missed opportunity; there are many systems thinking tools that could strengthen the development of both the Change- and the Action Model. The SAMEL approach to using these tools to develop programme theory is presented in the next chapter.

In the following section, I describe how I went about practically developing the content of the SAMEL Toolkit using bricolage.

SAMEL Pragmatist Framework: Bricolage

The SAMEL Toolkit's conceptual background is TBE and the systems concepts approach. The theory-based MEL components (see Table 7) constituted the skeleton of the Toolkit, and the systems concepts became the muscle. As described above, my approach to the systems concepts included finding appropriate systems thinking tools that would guide

the application of each concept in each part of the MEL development process. As such, the practical approach taken to developing the Toolkit's content and making decisions was framed by the pragmatic tradition of bricolage.

Anthropologist Claude Lévi-Strauss borrowed the French term *bricoleur*, referring to craftspeople who creatively and innovatively use spare materials to construct new crafts and artifacts, as a metaphor for making meaning (Lévi-Strauss, 1966). Some qualitative researchers and academics, such as Denzin and Lincoln (2005) and Kincheloe and Berry (2004), have since advocated the use of bricolage in research as an approach based in plurality, multiplicity, flexibility, and, importantly (in light of systems thinking), emergence (Rogers, 2015).

By embracing bricolage, the SAMEL Toolkit embodies both method and methodological pluralism — the use of insights from several methodological ideas and principles (i.e., branches of systems thinking), as well as the use of a range of methods and tools in support of these ideas (Boyd et al., 2004). The strength of pluralism lies in the ways in which it equips us with a diversity of perspectives that could aid flexibility and inclusivity when working in different contexts (Midgley, 2000). When working with a variety of methods and methodologies, we are able to view the work through multiple lenses, instead of having our vision narrowed by a single viewpoint. This grew in importance to me as I became more familiar with the concepts of systems thinking, particularly the emphasis placed on inclusivity and the importance of operating from multiple perspectives.

Bricolage is inherently interdisciplinary, and is thus vulnerable to the criticism of superficiality (Kincheloe, 2001) — covering a broad array of fields, but not deeply or comprehensively (i.e., 'jack of all trades, master of none'). This was a notable trade-off to be

made in the development of the Toolkit, particularly in light of the critiques of the systems concepts approach.

However, I argue that a pragmatic approach to systems thinking necessitates the use of bricolage, especially as it applies to MEL. Both MEL and systems thinking are vast fields that span several disciplinary boundaries. Integrating multiple perspectives (i.e., disciplines, fields, tools, and concepts) reflects the inherent nature of these fields, and thus strengthens the Toolkit by expanding the resources and concepts that could be applied to MEL. Such a dynamic, flexible, and multi-perspectival approach is inherently congruent with the field of systems thinking. Rigidly adhering to the boundaries of a single discipline and methodology is, I believe, contrary to the emergence and dynamism necessary for working within complex social contexts.

Making Theoretical Decisions: A Reflexive Pause

The process of integrating systems thinking tools and concepts with MEL procedures necessitated familiarising myself with as many accepted systems thinking disciplines as possible within a relatively short time frame, and navigating the literature was slow and challenging.

To try and bound the search and synthesis, I decided to focus my study on books, as they tend to offer historical overviews by pioneering authors in their respective fields, and have academic legitimacy. I looked for commonly cited volumes that (a) introduced systems thinking broadly (as opposed to discipline-specific approaches) and (b) provided an overview of a variety of methods that could be applied to evaluation. The volumes that I drew on most for the Toolkit were: *Thinking in systems: A primer* (Donella Meadows), *The fifth discipline* (Peter Senge), *Systems approaches to managing change* (Martin Reynolds and Sue Holwell), *Systems thinking for social change* (David Stroh), *Systems concepts in action* (Bob Williams

and Richard Hummelbrunner), *System concepts in evaluation* (Bob Williams), and *Developmental evaluation* (Michael Quinn Patton).

The works authored or co-authored by Bob Williams, an Australian evaluator, were particularly helpful in terms of their application of systems thinking to evaluation, as well as their introduction to a variety of systems methodologies and concepts. These books became the greatest inspiration for the Toolkit, particularly *Systems concepts in action*.

As a newcomer to systems thinking but an experienced MEL professional, developing the Toolkit was both stimulating and humbling. I thoroughly enjoyed studying the field and imagining creative ways to blend systems thinking and MEL, but realising the vastness of the field was intimidating, and provided much food for thought, especially with regard to rigorous methodological application.

During the Toolkit's design, I was aware of the tension between methodological rigour and the practicalities of MEL facilitation. Based on my professional experience, I had a good idea of what could reasonably be facilitated in a ToC workshop, for example, and what NPO teams would likely expect from a MEL process. I also knew what most funders would expect of the NPOs. Many of these knowns were incongruent with what I was reading in the systems thinking literature, and I had to continuously make trade-offs between the requirements of systems thinking methodologies and real-life practicalities.

Bolstered by my practical knowledge of MEL facilitation, all decisions centred on the systems concepts, rather than the methodologies. Instead of thinking of ways to equip evaluators with methodological knowledge, I decided to use the methodologies to bolster evaluators' understanding of the concepts.

As I continued to study methodologies, I mapped them onto the MEL phases I wanted the Toolkit to address: a situation analysis; a programme theory; a monitoring plan; an

evaluation plan; and a learning strategy. This was a time-consuming and iterative process, which took the better part of six months.

Chapter Conclusion

The SAMEL Toolkit is a workshop-based approach designed to help NPOs develop a MEL framework. It aims to introduce evaluators to systems thinking tools and concepts, providing guidance on practically facilitating these tools through a series of five participatory workshops. The Toolkit's design follows a structured outline, with each workshop targeting specific MEL components: a situation analysis, programme theory (ToC and ToA), monitoring framework, learning strategy, and an ideal evaluation plan. These aspects are described in detail in the following chapter.

The use of five participatory workshops addresses the need for less resource- and time-intensive approaches to systems-based MEL, such as DE or OM. By grounding the application of the Toolkit content in workshop facilitation, it fosters collaboration between evaluators and NPO teams, rather than focusing solely on capacitating the evaluator.

The Toolkit was deliberately designed to make systems thinking accessible to evaluators unfamiliar with the field who may feel overwhelmed or intimidated by anthologies or volumes that present a large number of tools to understand and choose from (e.g., USAID, 2020; Williams & Imam, 2006). This helps to address critiques of the systems concepts, such as their potential for misuse and varying interpretations, by providing specific instructions and pre-chosen tools for applying these concepts in MEL development. This pragmatic approach, drawing on bricolage, aims to make systems thinking accessible while still maintaining methodological rigour – thus blending minimum and maximum specifications of systems methodologies (Jackson, 2023b).

A component of this accessibility is the deliberate use of traditional and familiar MEL concepts and terminology, enabling widespread adoption among evaluators, NPOs and funders. This approach draws on key learnings from studies that identified OM's use of jargon and non-traditional processes as potential barriers to adoption (Smith et al., 2012).

While the Toolkit has examples particularly associated with the ECD sector (for purposes of the test), the general principles of each workshop can be adapted and applied to any context, as the systems thinking tools themselves are sector-neutral. A core argument made by this thesis, and the Toolkit, is that it can be applied in any social context, as social contexts are inherently complex. As such, the Toolkit avoids a contingency approach to its application, whereby it is only deemed appropriate for use among NPOs working in complex, uncertain, or unpredictable contexts – an approach taken by DE (Patton, 2011) and RMA (Mierlo, 2010).

The conceptual foundation of the SAMEL Toolkit is rooted in TBE, which aids in understanding programme effectiveness by examining stakeholders' assumptions about programme design and implementation. Specifically, it draws on Chen's (2015) Action Model/Change Model Schema, which emphasises the relationship between programme outcomes and implementation within an open programme system that is influenced by external environments.

While Chen's (2015) model is comprehensive, it lacks integration with systems thinking concepts, which I believe can strengthen the development of both change- and action models. The SAMEL Toolkit addresses this by incorporating systems concepts and related tools into the development of a theory-based MEL framework (described in the following chapter).

The use of TBE was a deliberate decision to support the Toolkit's adoption, recognising that while social change is inherently complex, emergent and unpredictable, programme theories are often required by funders of South African NPOs. Furthermore, while approaches like OM or DE emphasise emergent change and adaptive strategies, they may overlook opportunities for empirical and predictive modelling of change. A systems-based programme theory thus provides an opportunity to leverage existing empirical evidence and experience while incorporating important contextual and systemic insights.

In the following chapter, I describe the content of the Toolkit in detail, including the theoretical underpinnings, as well as the evaluation experts' feedback and how it was integrated into the Toolkit's design.

Chapter Seven: Content of the SAMEL Toolkit

The following sections detail the content of each chapter of the SAMEL Toolkit, focusing on a description of the systems thinking tools that I chose to include and the rationale for doing so. I also describe the SAMEL workshops and how the Toolkit content is incorporated into each workshop.

For the sake of brevity, this chapter focuses on the final version of the Toolkit that was tested, as opposed to the version that was reviewed by experts. This is to give the reader a complete understanding of the Toolkit that was tested and, thus, the study's findings. However, the changes made to the Toolkit following the expert review are noted throughout.

Links^{viii} are provided to an online version of the SAMEL Toolkit throughout this chapter, should the reader be interested in reading further or consulting particular resources.

In addition to an introductory chapter, the Toolkit consists of four chapters that detail its four phases of MEL development: Scoping, Focusing, Monitoring & Learning, and Evaluation. Each Toolkit chapter is structured in the same way, illustrated in Figure 12.

^{viii} Links to the Toolkit have been removed from this copy of the thesis, as the SAMEL Toolkit Intellectual Property belongs to the University of Cape Town.

Figure 12

SAMEL Toolkit Chapter Structure

1. Chapter objectives, learning outcomes, task summary and workshop outputs

2. Chapter introduction and rationale to MEL component in focus

3. Systems thinking tools for the development of the MEL component

4. Facilitation guidelines for the associated workshop

5. Worksheets and resources

THEORY OF CHANGE CHECKLIST

WHAT MAKES FOR A POOR TOC?	CHECKPOINT	CHECKLIST
Emphasis on planning as opposed to understanding the programme logic - why does it work?	The ToC adequately communicates how the programme intends to achieve its desired change, and why.	<input type="checkbox"/> The relationships between all ToC components are communicated clearly. <input type="checkbox"/> Impact is clear.
Rigid formats are unable to accommodate ongoing and adaptive change.	The programme team understands the need to continuously reflect and update the ToC, and have the capacities to do so.	<input type="checkbox"/> Team members understand that the ToC should be reflected on and updated periodically. <input type="checkbox"/> Team members feel a sense of ownership over the ToC. <input type="checkbox"/> Team members feel capable of amending it as needed.
Emphasis on filling in boxes as opposed to in-depth engagement with systems principles.	The ToC reflects a development process in which systems and boundaries, relationships, and perspectives were taken into account using systems thinking tools and concepts.	The ToC development process included deliberate and thoughtful consideration of: <ul style="list-style-type: none"> <input type="checkbox"/> Systems and boundaries <input type="checkbox"/> Relationships <input type="checkbox"/> Perspectives
Ignoring interactions between outcomes and	The ToC incorporates the factors believed to influence	<input type="checkbox"/> Causal mechanisms are included in the ToC.

Expert Review

An expert review of the Toolkit was conducted by three seasoned evaluators prior to it being field-tested. The methodology of the review was described in Chapter Five.

The review was conducted to assess the Toolkit's effectiveness and utility from the perspective of professionals in the field of evaluation and systems thinking. This feedback was integral to refining the toolkit and ensuring it would meet the needs of evaluators and NPOs. The experts provided overall feedback, as well as some more focused suggestions and critiques, which, for the sake of brevity and flow, were integrated into the description of the Toolkit in the following pages. The experts' overall feedback is provided at the end of this chapter.

Toolkit Chapter 1: Introduction

The introductory chapter of the SAMEL Toolkit provides an overview of and rationale for the Toolkit's systems thinking approach. It briefly introduces systems thinking and systems concepts, explains the relevance of systems thinking to MEL design, and then introduces the Toolkit (see Figure 13 for an excerpt). The expert reviewers suggested no changes to this chapter.

As shown in Table 7, the Toolkit's SAMEL process consists of four chapters (or steps):

1. Scoping (Workshop 1);
2. Focusing (Workshop 2: Theory of change and Workshop 3: Theory of Action);
3. Monitoring & Learning (Workshop 4); and
4. Evaluation (Workshop 5). These are discussed in the subsequent sections.

Figure 13

Excerpt from Chapter 1 Explaining Systems Thinking Rationale

WHY SYSTEMS THINKING?

In a seminal book on systems thinking and organisational management, Peter Senge wrote:

The essence of the discipline of systems thinking lies in a shift of mind: seeing interrelationships rather than linear cause-effect chains, and seeing processes of change rather than snapshots^{11(p. 92)}

This quote highlights why systems thinking is an especially useful lens for MEL, which, at times, can be too linear or reductionist. Theories of change, logical frameworks and randomised control trials typically depict programme pathways as a linear causal sequence, beginning with the programme inputs and activities, and ending in participant outcomes and impact¹². Social problems, however, are rarely stable and linear; they are adaptive and interconnected, and they are subject to multiple differing perspectives on what they are, and how to solve them. Authors^{13,14} have urged civil society, including Non-Profit Organisations (NPOs) and funders, to relinquish mechanistic and linear approaches to strategy, and replace them with a more adaptive learning strategy that can better respond to the uncertainty and emergence that characterise complex social problems.

<div style="background-color: #0070C0; color: white; border-radius: 15px; padding: 5px; margin-bottom: 5px;">CONVENTIONAL / LINEAR THINKING</div> <div style="background-color: #ADD8E6; border-radius: 15px; padding: 5px; margin-bottom: 5px; text-align: center;">The causes of problems are easy to identify and trace</div> <div style="background-color: #ADD8E6; border-radius: 15px; padding: 5px; margin-bottom: 5px; text-align: center;">Interventions and policies that achieve short-term success will also achieve long-term success</div> <div style="background-color: #ADD8E6; border-radius: 15px; padding: 5px; text-align: center;">To optimise system performance, we must optimise the performance of its parts</div>	VS	<div style="background-color: #0070C0; color: white; border-radius: 15px; padding: 5px; margin-bottom: 5px;">SYSTEMS THINKING</div> <div style="background-color: #ADD8E6; border-radius: 15px; padding: 5px; margin-bottom: 5px; text-align: center;">The relationship between problems and their causes may be indirect or obscured by other factors</div> <div style="background-color: #ADD8E6; border-radius: 15px; padding: 5px; margin-bottom: 5px; text-align: center;">Quick fixes, or short-term symptom alleviation, can have unintended consequences and undermine long-term success</div> <div style="background-color: #ADD8E6; border-radius: 15px; padding: 5px; text-align: center;">To optimise system performance, we must improve the relationships between the system parts</div>
--	----	---

Note. Adapted from *Systems thinking for social change* (p. 19), D. P. Stroh, 2015, Chelsea Green Publishing. Copyright 2015 by David Peter Stroh.

Toolkit Chapter 2: Scoping

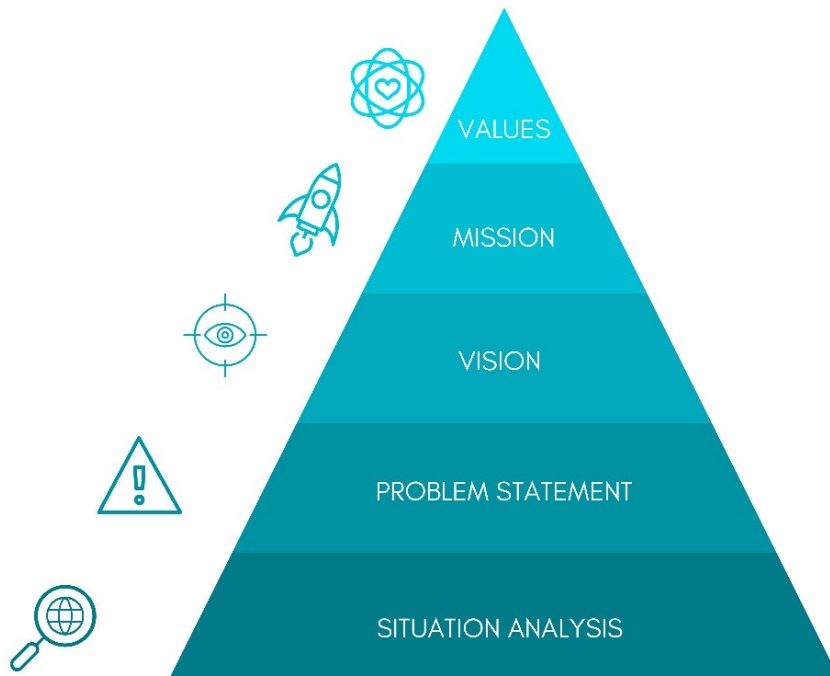
While borrowing from Chen’s (2015) reference to *programme scope* as describing a programme’s goals and focus, the Toolkit’s version of the programme scope is aligned with Funnel and Rogers’s (2011) first two steps in programme theory development: conduct a situation analysis to understand the problem (and opportunities) of interest, and decide on

the aspects of the problem on which the programme will focus. Chapter Two of the Toolkit thus outlines the development of the programme scope, which consists of a problem statement, vision, mission, and values. The inclusion of a vision and mission statement was inspired by the early stages of Outcome Mapping (developed by Earl et al., 2001).

The SAMEL Toolkit bases the programme scope on a brief situation analysis conducted by the evaluator. The rest of the programme scope is then built upon the situation analysis with the NPO team (see Figure 14). Using the information from the situation analysis, a clear and concise problem statement is defined. The problem statement then informs the programme vision, which, in turn, informs the programme mission and values.

Figure 14

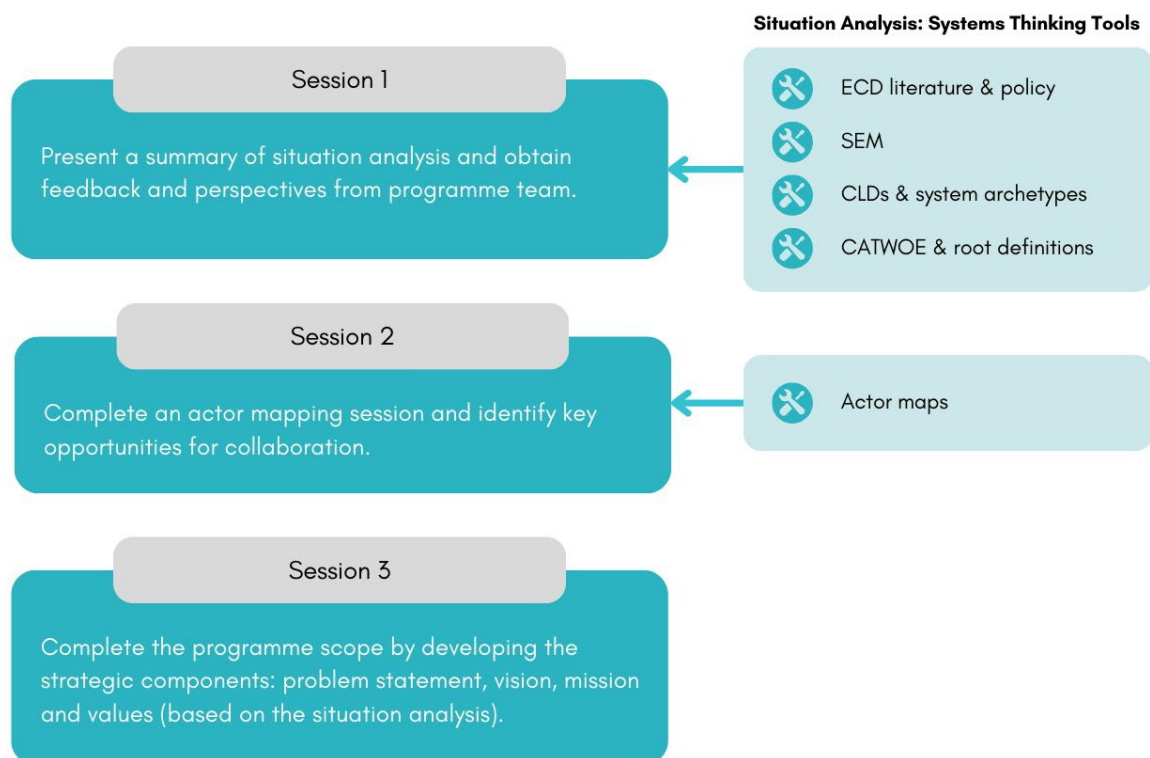
SAMEL Constituents of Programme Scope



The SAMEL Scoping step’s workshop comprises three sessions to develop the programme scope. In the first session, the evaluator presents a draft situation analysis to the programme team for review, reflection, input, and discussion. The draft situation analysis is prepared prior to the workshop using a number of systems thinking tools introduced by the Toolkit (described below). In the second session, the team develops an actor map to deliberate on relationships that are relevant to the programme and to identify opportunities for collaboration. In the third and final session, the programme scope is completed using the situation analysis. Figure 15 provides an overview of the sessions of the Scoping workshop.

Figure 15

Scoping Workshop: Summary of Sessions



Evaluator Preparation for Scoping Workshop: Situation Analysis

Conducting a situation analysis involves gathering information from a wide variety of sources regarding the context or situation within which the programme is (or will be) situated. This includes, amongst other things: the history of the problem that needs to be addressed, the groups of people and the systems involved, the root causes and consequences of the problem, and the perceptions of those involved (Funnel & Rogers, 2011).

An important quality of a situation is that it is not a concrete, taken-for-granted reality, but is co-constructed by various people who may share a particular perspective. As Checkland and Poulter (2010, p. 201) explain:

Nothing is intrinsically 'a situation'; it is our perceptions which create them as such, and in doing that we know that they are not static; their boundaries and their content will change over time.

Thus, it can be seen as a snapshot of reality at a particular point in time, shaped by those who view it. Developing a programme theory (in the next SAMEL chapter and workshop) grounded in a situation analysis thus ensures that the programme design is responsive, relevant, and representative of multiple perspectives of the problem. The SAMEL situation analysis is based on an amended list of questions from Funnel and Rogers's (2011) book *Purposeful program theory*, broken down according to the systems concepts.

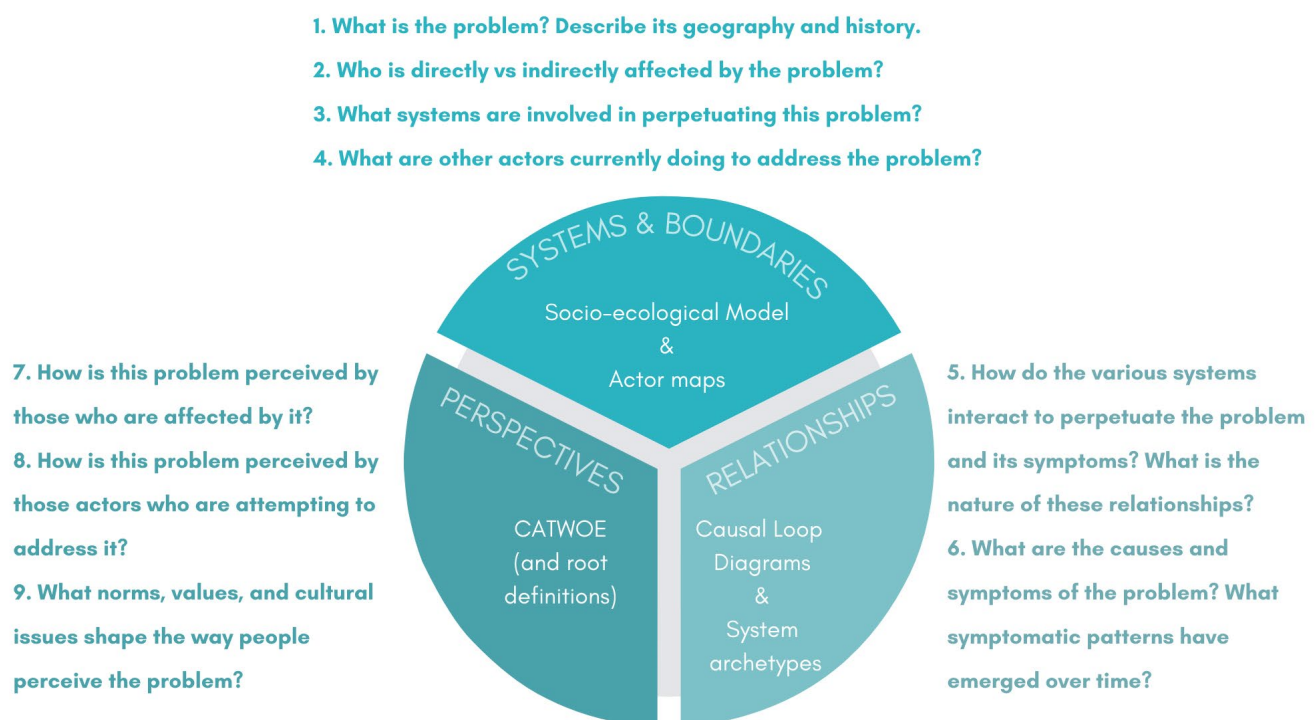
While researching the situation analysis, the Toolkit advises evaluators to first consult literature to gain an understanding of the policy implementation context, and then to critically deliberate on the sources to consult. A library of South African ECD resources is provided to help identify key literature, and a reflection worksheet is provided to prompt evaluators to critically consider the voices they are choosing to elevate in their analysis, which ones they are marginalising, and why.

To prepare for the Scoping workshop, evaluators are advised to develop draft answers to the situation analysis questions, which are then presented in Session 1 of the workshop. This is done using a number of systems thinking tools introduced in the Toolkit, described in the sections below: the SEM (Bronfenbrenner, 1979), causal loop diagrams and system archetypes (Senge, 1990), and CATWOE from Soft Systems Methodology (SSM; Checkland, 1981).

Causal loop diagrams, system archetypes, and SSM were first identified from the book *Systems concepts in action* by Bob Williams and Richard Hummelbrunner (2011), and then matched to the amended list of Funnel and Rogers’s (2011) situation analysis questions. Figure 16 summarises the situation analysis questions, alongside the systems concepts and associated tools proposed to answer the questions. These are described in the subsequent sections.

Figure 16

SAMEL Situation Analysis Questions with Associated Systems Concepts and Tools



Systems and Boundaries: The SEM. The first four questions look at the system ‘parts’ of the problem (the individuals involved) and the boundaries of the situation. The Toolkit proposes the use of Bronfenbrenner’s (1979) SEM (introduced in Chapter Four) and actor maps to help answer these questions (see Figure 16).

The SEM was developed by Urie Bronfenbrenner as a means to understand the complexity of human development by placing an individual within a set of nested, interacting systems (Bronfenbrenner, 1979; Hayes et al., 2017). The model has been widely used in the context of child development, as Bronfenbrenner was outspoken about the role of society in successfully rearing children (Hayes et al., 2017). The model was described in depth in Chapter Four.

Systems are not real-world entities, but conceptual constructs that help us break down, understand, and describe complex realities (Ulrich & Reynolds, 2010). As such, the SEM (Bronfenbrenner, 1979) provides an ECD-specific heuristic to assist evaluators and programme teams in identifying systems of interest. Evaluators using the SAMEL Toolkit are encouraged to use the SEM (Bronfenbrenner, 1979) to identify systems of interest. A SEM worksheet is provided to assist evaluators in this process.

Systems and Boundaries: Actor Maps. To understand the actors (or system parts) currently involved in the situation, the Scoping workshop, in Session 2, includes an actor mapping session. This is the only situation analysis component that the Toolkit proposes be implemented in the Scoping workshop, as it is considered an important planning tool.

Actor maps are a visual depiction of the key organisations or individuals that make up a given system, as well as the relationships between them (Gopalakrishnan & Clarke, 2015). This exercise extends the SEM (Bronfenbrenner, 1979) by aiding the identification of

particular systems or actors of interest in the situation (system parts in the socio-ecological system) and how they are related, as well as the fleshing out of particular intervention strategies that may depend on partnerships and collaboration (outside of the programme's boundaries).

Actor mapping guidelines developed by Gopalakrishnan and Clarke (2015) for the Foundation Strategy Group are provided in the Toolkit to guide evaluators through this session. Evaluators are advised to use South Africa's ECD Essential Package (Ilifa Labantwana, 2014) as a skeleton for the map, which divides actors into five key domains for holistic child development, aligned to South Africa's National Integrated ECD Policy (RSA, 2015): nutritional support, social services, support for primary caregivers, stimulation of early learning, and maternal and child health. Evaluators are given an actor map template to practise before this session.

Relationships: Causal Loop Diagrams and System Archetypes. The fifth and sixth situation analysis questions concern relationships and patterns that characterise the situation of interest (see Figure 16). Two types of relationships are deemed relevant during Scoping: (1) relationships between actors in the system and (2) relationships between systemic variables that are creating problematic social patterns. The former type is dealt with in the actor mapping session, while the Toolkit suggests causal loop diagrams (CLDs) and system archetypes as tools to identify cause–effect relationships and problematic patterns. This forms part of the situation analysis in Session 1 of the workshop.

The Scoping chapter differentiates between aspects of a situation that are easily observable, such as incidence rates (problem symptoms), versus structural elements of the situation that are, over time, reinforcing problematic patterns in the situation of interest. Identifying causal loops, vicious cycles, and system archetypes is suggested as a means to

unpack these structures and resultant patterns, helping the NPO team to gain an in-depth understanding of what is causing problem symptoms in the situation (and thus what to target through intervention).

In 1990, engineer and influential systems scientist Peter Senge introduced the notion of a learning organisation in his seminal book *The fifth discipline: The art and practice of the learning organization*. In order to build organisations that can learn and adapt to achieve effectiveness, Senge argued, systems thinking helps us to clearly observe ongoing patterns and determine how to change them (Senge, 1990, 2006). While using the general term *systems thinking*, Senge drew heavily on system dynamics, a branch of the systems field concerned with the dynamic behaviour of complex systems (Sterman, 2001). Fellow engineer turned management scientist Jay Forrester founded the field of systemic dynamics in the 1960s to understand and model the dynamics of social organisations (Forrester, 2007a). It is grounded in the theory of non-linear dynamics and feedback control developed in mathematics, physics, and engineering (Sterman, 2001). Senge draws on key system dynamics concepts to unpack the system interrelationships and structure that underlie dynamic patterns of system behaviour in organisations. Similarly, in his book *Systems thinking for social change*, David Stroh (2015) examines the same concepts in pursuit of understanding how social change is created by policy systems and social interventions.

A major advantage of system dynamics is uncovering the systemic processes and structures that create consistent, long-term patterns of behaviour, instead of focusing on surface-level events or crises (Stroh, 2015). This is the difference between asking: *What happened that we need to react to?* (e.g., an employee strike) versus *Why did this happen?* (e.g., poor working conditions). Asking *why* questions helps us take a step back and consider the root causes of system patterns or trends (Stroh, 2015).

CLDs have their foundation in the field of cybernetics, which is primarily concerned with feedback within a given system (Ashby, 1956; Beer, 1972; Weiner, 1948). CLDs illustrate behavioural patterns in a system by depicting circular cause-and-effect relationships between variables, known as feedback loops (Meadows, 2009; Williams & Hummelbrunner, 2011).

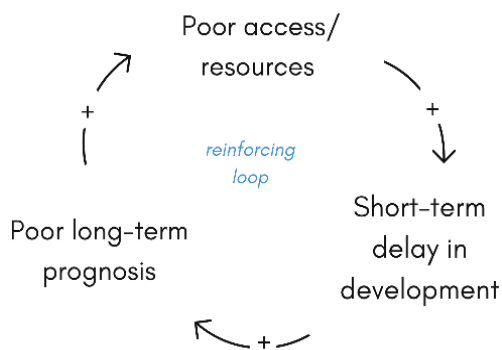
Some feedback loops are reinforcing, where variables respond to each other in the same direction, creating exponential growth (Meadows, 2009; Stroh, 2015; Williams & Hummelbrunner, 2011). These relationships are depicted using '+' in CLD diagrams (see below). For example, the more money you put into a savings account, the greater the interest earnings will be. Feedback loops can also be balancing or corrective; as one variable increases, the other lessens (Meadows, 2009; Stroh, 2015; Williams & Hummelbrunner, 2011). These are illustrated using '-' in CLD diagrams. For example, when you withdraw money from your savings account, the interest rate decreases.

While CLDs are useful for illustrating system behaviour, it can be difficult to know where to start. The Toolkit thus proposes vicious cycles and system archetypes as baseline heuristics to help evaluators and programme teams identify problematic patterns.

Vicious cycles involve reinforcing feedback loops that multiply and snowball, causing "runaway destruction" (Meadows, 2009, p. 31). As one variable increases, so does the other (see Figure 17 for an example). While reinforcing feedback can be positive, if left unchecked, can cause system malfunction or breakdown (Meadows, 2009).

Figure 17

Example of a Vicious Cycle



Per the example in Figure 17, it is well established that early learning provision in South Africa reflects socio-economic inequality. Young children from lower income groups and rural regions are less likely to have access to an early learning programme (Ashley-Cooper et al., 2019; Hall et al., 2019), because caregivers cannot afford the fees, or because there are no services available in their area. This has profound consequences for children’s development. Children of parents in the lowest income bracket demonstrate poor developmental outcomes in comparison to their counterparts who have wealthier parents (Snelling et al., 2019a), and thus enter formal schooling with delayed development and poor school readiness, which are then exacerbated by dysfunctional, low-resource schooling (Ashley-Cooper et al., 2019). These, in turn, amplify their learning deficits, which ultimately hamper their future success and access to socio-economic resources.

System archetypes are generic CLDs that represent combinations of reinforcing and balancing feedback that produce commonly observed patterns of system behaviour (Senge, 1990). The archetypes tell systems stories that can be applied to any system or situation. Using these archetypes as templates can guide thinking about, and aid identification of, the non-linear relationships that may be causing problematic patterns in the situation of interest,

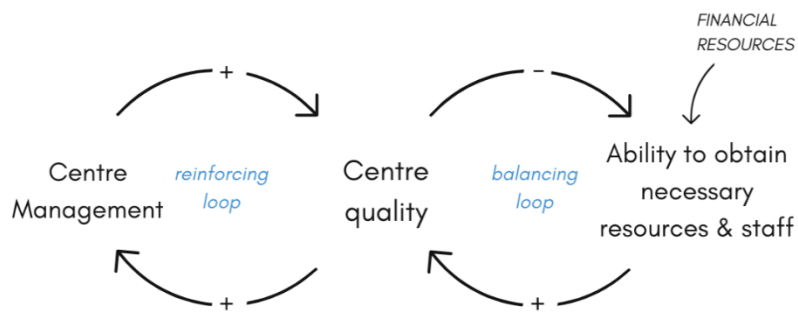
instead of creating unique CLDs from scratch. The Scoping chapter provides several examples of system archetypes that might be relevant for an ECD programme, such as the Limits to Growth archetype.

This archetype describes a situation in which a system's performance gradually decreases due to a particular constraint or limiting force (Senge, 1990). It consists of a reinforcing loop that illustrates system growth due to a particular course of action. The more efforts are applied, the more the system's performance increases. However, as the system continues to grow, it eventually reaches a plateau or begins to decline as it comes up against a constraint. Here, a balancing loop depicts how a constraint creates a slowing action that decreases performance.

An example of this archetype can be seen in fee-based ECD centres (see Figure 18). A South African study by Biersteker et al. (2016) on the quality of ECD centres found that centre quality was most strongly predicted by the fees charged and the quality of the centre's management. Centres in the most poverty-stricken of areas were of the lowest quality. The amount of money raised by a centre can be transformed into quality through the ability to employ and retain skilled staff, purchase learning materials, and develop facilities and infrastructure (Biersteker et al., 2016). Thus, centre quality cannot be increased by improving only the management of the centre; financial resources will always play a central role in the centre's quality.

Figure 18

Limits to Growth Example



Perspectives: Soft Systems Methodology and CATWOE. The final three situation analysis questions address the ways in which the problem of interest is perceived by relevant actors, including the programme team. CATWOE, a component of SSM, is suggested as a conceptual process to help the programme team understand the various ways in which different actors in the system perceive the problem, as part of the situation analysis discussion (Session 1 of the workshop).

SSM was first developed by Peter Checkland (1981) as an alternative to the use of 'hard' systems engineering methods in addressing management problems (Checkland, 2000). As described in Chapter Three of this thesis, 'hard' methodologies see the world as comprised of real-world systems that require standardised intervention, while 'soft' methodologies acknowledge that the world looks different to everyone. Management- and social problems are characterised by complex relationships and interactions between people, which originate from different perspectives and worldviews (Checkland & Poulter, 2010). The cornerstone of the methodology is the recognition of how these different worldviews lead to different perceptions of the situation at hand, including the most effective way to intervene and improvement it (Checkland 1981; Checkland & Poulter, 2010). This is important for a situation

analysis, as the team deliberately considers how other key stakeholders perceive the situation, and what they perceive as possible interventions (in addition to the team's chosen interventions).

SSM comprises several steps that make up an iterative learning cycle designed to uncover these different worldviews, with a view to taking action that would change the situation (Checkland 1981; Checkland & Poulter, 2010). The learning cycle begins with a particular problematic^{ix} situation that has two major characteristics: the situation contains people trying to act purposefully and with intent, and the situation is perceived differently by people with various worldviews. One way to understand this situation is to use rich pictures — completely unstructured, free-hand drawings aimed at including as many aspects of the situation as possible. Other useful concepts here are the relevant roles, norms, values, and power dynamics that characterise the situation. The next step in SSM is to identify the purposeful activity systems that are active within the situation.

Purposeful activity systems consist of a transformation process and a worldview, whereby a group of actors (e.g., an NPO) will perform a set of activities aimed at transforming the state or conditions of another group of actors (e.g., programme participants). According to SSM, these activity systems are based on a particular worldview; thus, multiple different systems can operate from multiple worldviews.

For example, a school may take the view that poor Grade 3 learning outcomes are due to poor-quality teaching. The school may then arrange that their teachers embark on

^{ix} Checkland and Poulter (2006, 2010) prefer the term *problematical* to *problem* as a way to capture the idea that a situation is perceived as problematic, rather than the assumption that there is an inherent problem with a clear-cut associated solution.

professional development courses aimed at transforming their instructional skills. An NPO may attribute poor learning outcomes to a lack of educational support at home, and direct efforts towards engaging caregivers in learners' education. Thus, each worldview determines the chosen transformation process.

In order to flesh out these purposeful activity systems, SSM offers a mnemonic device, CATWOE, which stands for Customers, Actors, Transformation, Worldviews, Owners, and Environment.

- Customers are the people who will either benefit from or be victimised by the transformative activities (e.g., programme participants).
- Actors are the people who will perform the activities that constitute the transformation (e.g., programme designers and implementers).
- Transformation is the process through which the customers will shift from one condition to another.
- Worldview is the assumptions and perspectives that determine how the actors see the world.
- Owners are the people or groups who can be regarded as owning the transformation process, and thus have the power to hamper or enable it.
- Environment is the external constraints that affect the transformation process, such as rules, norms, practices, resources, etc.

The Scoping chapter guides evaluators to use CATWOE to develop root definitions as part of the situation analysis, preceded by a stakeholder analysis (worksheet provided), in which the programme team identifies key stakeholders whose perspectives should be considered. The root definition concisely defines the activity system's worldview, what it does,

how, and why, using the information from CATWOE (see Figure 19). The term ‘root’ is used here to suggest that this is one of many possible definitions of the system, as each definition reflects a single worldview (Checkland & Poulter, 2010). Evaluators are provided with an SSM worksheet to practise CATWOE and root definitions.

Figure 19

Example of Root Definitions using CATWOE

Root definitions using CATWOE elements	
1	A school-owned system (O) which contracts professional development staff (A) to improve teachers' (C) instructional practices (T) in order to improve Grade 3 learning outcomes (W).
2	An NPO-owned system (O) whose professional staff (A) provides educational training to caregivers (C) in order to improve their engagement with their children's education (T) and thus improve their children's learning outcomes (W).

The biggest advantage of SSM is the systematic appreciation for, and consideration of, different perspectives that are at play in any given social situation. Applying CATWOE and root definitions to a situation analysis helps to ensure that the analysis is not conducted from a single worldview, but that it richly captures the multiple interacting perspectives that are important to understand.

This may assist the programme team in two major ways: (1) identification of potential partners with whom the team can collaborate; and (2) comparison of the team's perspectives and worldviews with those of others and identifying whether there is an opportunity to learn from others.

Scoping Summary

The Scoping chapter of the SAMEL Toolkit establishes a MEL foundation by assisting the evaluator and NPO team in developing an in-depth understanding of the situation to

which their programme is intended to respond. Prior to the Scoping workshop, the Toolkit guides the evaluator to develop a draft situation analysis bolstered by systems thinking tools, including the SEM, CLDs, system archetypes, and CATWOE. These tools are intended to broaden both the evaluator's and the NPO team's understanding of the systems involved in the situation of interest, the structural causes of problematic patterns and trends, and other perspectives on the situation. These insights are presented to the NPO team in the first session of the Scoping workshop for discussion and feedback.

In the second session of the workshop, the evaluator guides the team to develop an actor map to visualise the key organisations and individuals involved in the situation, as well as the relationships between them. This exercise highlights important stakeholders in the situation of interest, thus aiding identification of potential opportunities for influence and collaboration.

The final session of the workshop is dedicated to completing the programme scope, which includes defining a clear problem statement, vision, mission, and values (as illustrated in the pyramid in Figure 14). This process is informed by the situation analysis and actor mapping, ensuring that the programme's strategic direction is grounded in a well-rounded understanding of the situation of interest.

Expert Feedback on Scoping Chapter

All three expert reviewers considered this chapter particularly useful and relevant to NPO teams, and believed that the tools discussed in this chapter would expand evaluators' systems thinking competencies. The two local reviewers, Allette and Naledi, made some specific suggestions for improvement, which were incorporated into the final Toolkit's design:

1. Originally, the Toolkit drew on an organisational problem statement, vision, mission, and values. Alletta suggested that the focus be on the programme.

2. Worksheet instructions were expanded and made clearer.
3. Links between the systems thinking tools and how they should be used to build on each other were made clearer (for the evaluator) in each Toolkit chapter.
4. Additional summary graphics were created.
5. Additional trigger questions were incorporated into the workshop guidelines to ensure greater clarity of concepts.

Toolkit Chapter 3: Focusing

The key objectives of the Focusing step are to develop the programme theory, which consists of a ToC and a Theory of Action (ToA), taking a systemic perspective on programme planning. In the Focusing chapter, a full workshop is dedicated to developing each theory.

A ToC explicitly states the process through which intended change is understood to occur for a programme's participants and target population (Chen, 2015; Funnel & Rogers, 2011). The ToA outlines how the programme will go about activating this change process (Chen, 2015; Funnel & Rogers, 2011). Referred to collectively here as *programme theory*, these two outputs constitute the programme's main focus and complete the programme's core strategy (in addition to the problem statement, vision, mission, and values that were established during Scoping).

The Focusing step draws significantly on the work done in the previous step (Scoping). While the Scoping step provides broad information on *what is*, the Focusing step hones in on *what to do*.

Theory of Change

The SAMEL approach proposes the use of interventions, context, mechanisms, feedback, outcomes, and impact as the main constituents of the ToC. These components

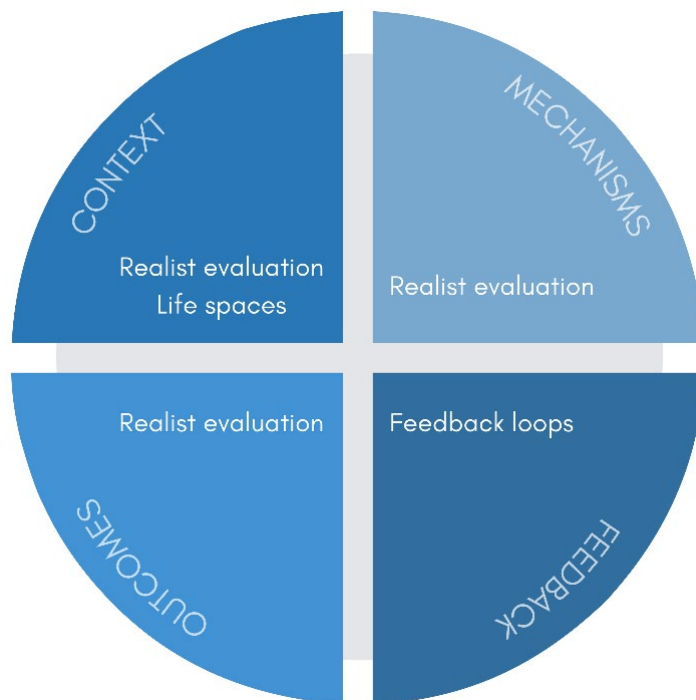
were inspired by Funnel and Rogers's (2011) approach to ToC development, with the addition of feedback and realist evaluation (discussed below). Evaluators are encouraged to visually depict the ToC elements in a diagram, with an accompanying descriptive narrative:

1. Interventions: The specific interventions (or sub-programmes) that are implemented by the programme team with the programme participants to initiate the intended change;
2. Context: A brief description of the context in which the programme will be implemented and the participant change that is expected to occur;
3. Mechanisms: Changes in the participants' reasoning and perceived resources that will determine how they respond to and engage with the programme, and, as a result, change.
4. Outcomes: The direct and observable results of mechanisms — how programme participants enact the inner transformation that has taken place as a result of participating in the programme. Outcomes represent the observable change in participants' knowledge, skills, behaviour, etc. that the programme intends to produce. If relevant, outcomes can also be developed for a secondary group of intended beneficiaries that may benefit from changes that take place amongst the primary participants (e.g., a programme that trains ECD practitioners as the primary participants, but intends change to take place amongst the practitioner's child attendees);
5. Impact: Long-term change that occurs at the community- or societal level after the programme has been established and running for an extended period; and
6. Feedback: Looping relationships between elements of the ToC that depict balancing or reinforcing feedback within the programme system.

Three systems thinking tools are provided to assist in the development of the programme's mechanisms, outcomes, context, and feedback. The elements of the ToC and the associated system thinking tools are illustrated in Figure 20.

Figure 20

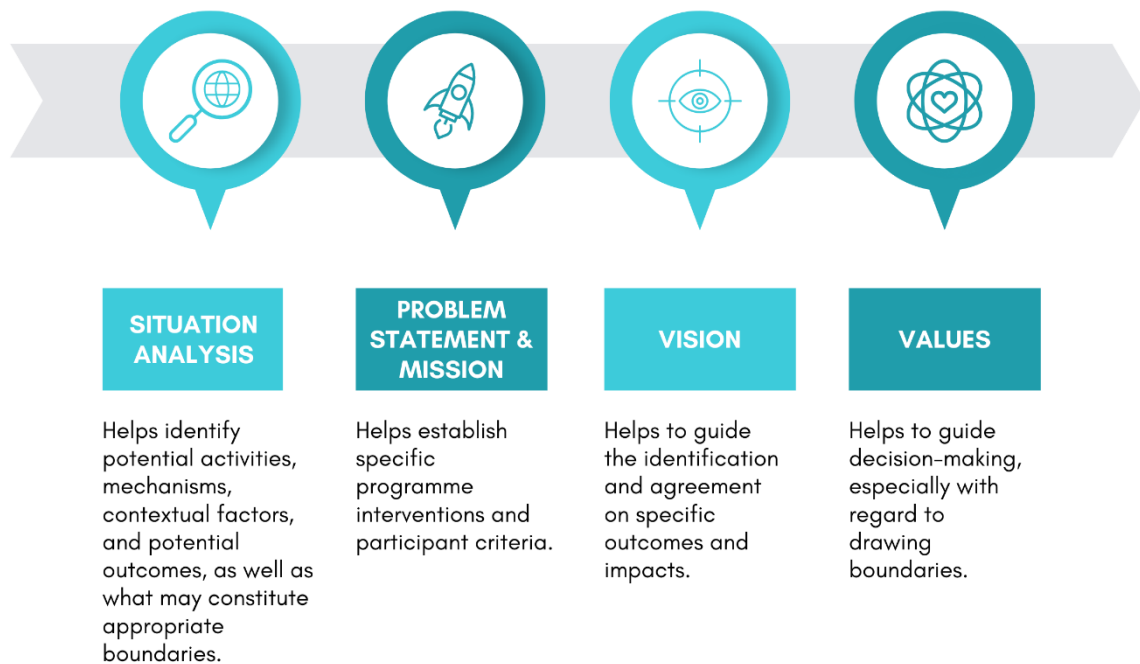
ToC Elements and Associated System Thinking Tools



Once the Scoping workshop has been completed, the programme team will be equipped with important contextual information (together with strategic direction) with which to make decisions about the programme's focus, or to revisit past decisions (see Figure 21). This is critical to ensure that the ToC is based on solid evidence, and not on selective issues or team assumptions that derail decision-making regarding the problem to be addressed or suitable solutions (Valters, 2014). A thorough engagement with the context also helps to ensure that the ToC outcomes are realistic and relevant to the programme and its context (Ringhofer & Kohlweg, 2019; Wilkinson et al., 2021).

Figure 21

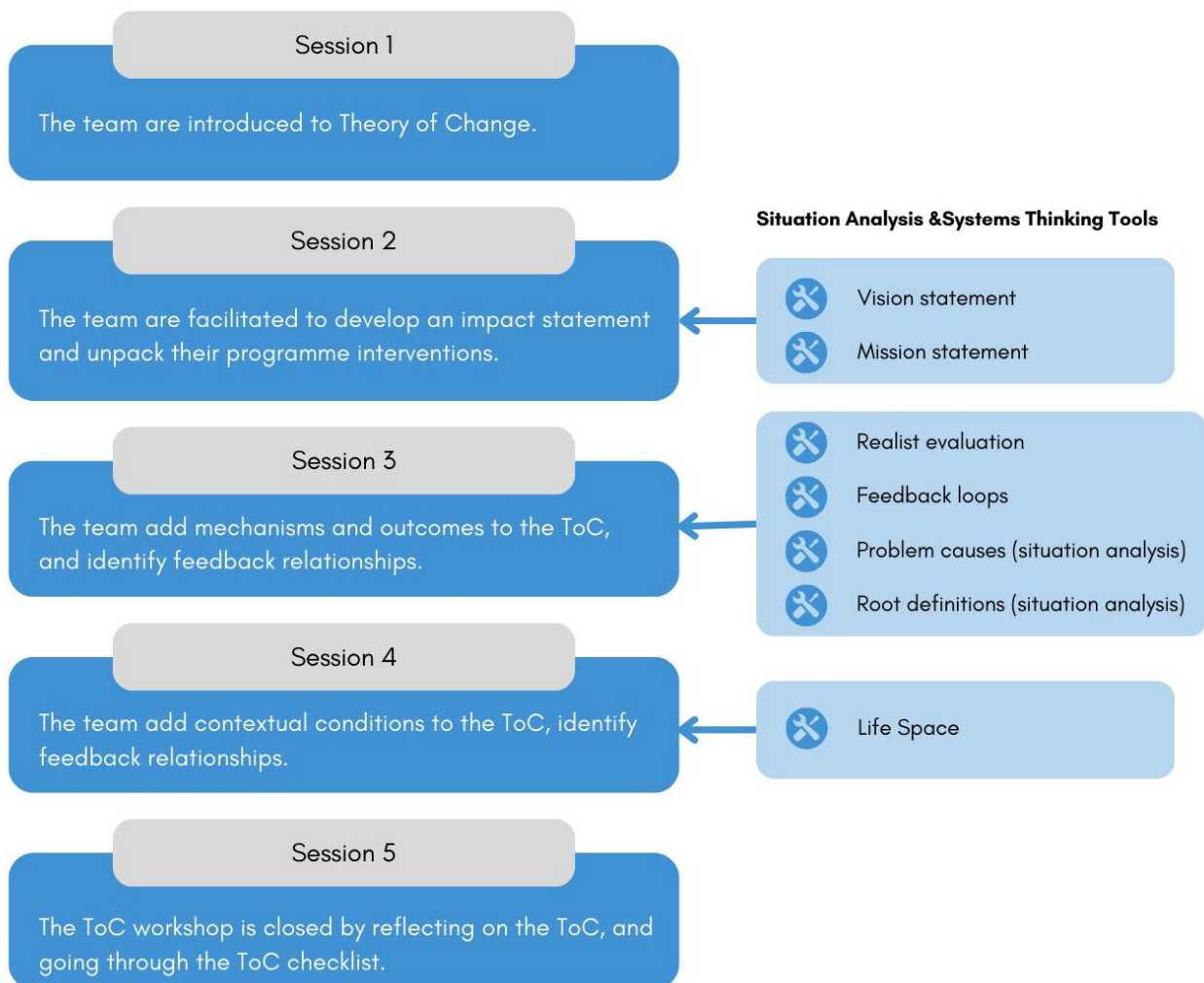
Function of Scoping Elements in ToC development



Prior to the ToC workshop, evaluators are advised to collate information from the situation analysis that could inform development of the ToC development process. The Toolkit also advises evaluators refer to the stakeholder analysis completed during the SSM step of Scoping, and to survey key stakeholders on the mechanisms, outcomes, and contextual conditions they believe are relevant to the ToC. This is done to avoid too many participants in a ToC workshop, while ensuring that other perspectives are included in the process. A summary of the ToC workshop sessions is provided in Figure 22.

Figure 22

ToC Workshop: Summary of Sessions



As illustrated above, the final session of the workshop includes a reflection on the ToC. The Toolkit provides a ToC checklist to help evaluators and programme teams assess the quality of the ToC. The checklist incorporates common critiques of ToCs (particularly linear models), and then provides checkpoints that help evaluators avoid pitfalls that lead to poor-quality ToCs.

In the sections that follow, I describe the systems thinking tools that the Toolkit provides in order for evaluators to support NPO teams to develop the ToC components above (in sessions three and four of the ToC workshop).

Realist Evaluation. The SAMEL approach assumes that mechanisms and outcomes are knowable in advance, meaning that the programme team has some idea of the change they expect to trigger through their actions, based primarily on empirical evidence, the findings of the situation analysis, and the programme’s vision and mission. However, evaluators are encouraged to remain aware of the fluid and emergent nature of social reality, and to communicate to teams that the ToC is never complete — it is a ‘snapshot’ taken at a particular time, and is subject to change. As such, the ToC should always remain a working document that is regularly updated as programme data are collected and reflected upon. This way, teams can remain attentive to change in the social environment and attuned to unintended and emergent outcomes.

Unpacking the causal pathways that underlie a programme’s intended impact is the central task in ToC development (Funnel & Rogers, 2011; Patton, 2012). While this necessitates the logical sequencing of programme interventions and linking these to programme outcomes, it is important to first consider the factors that inform these relationships. This argument has been put forth by many evaluation scholars (e.g., Chen, 2015; Funnel & Rogers, 2011; Pawson & Tilley, 1997), but it is not universally applied to theory development, especially in the case of linear models like logframes and logic models. All explanatory variables are typically put into the column labelled ‘*Assumptions*’, without much consideration of their precise nature and the importance of these factors to programme success.

The philosophy of realism, while not strictly a systems methodology, is a particularly useful lens through which to understand causal pathways. The SAMEL Toolkit attempts to workshop these pathways by using realist evaluation, life space modelling, and feedback loops to enhance teams’ understanding of the relevant causal pathways.

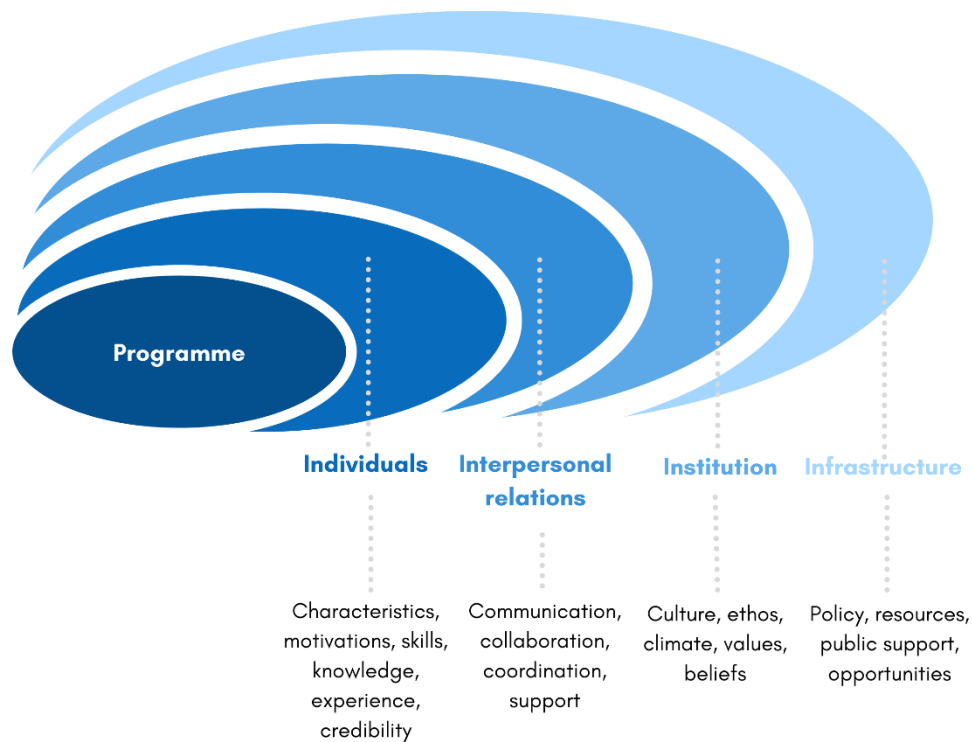
Realism takes the stance that reality or a real world (comprising both concrete and abstract phenomena) exists independently of our perception of it, and that, to understand reality, we need to move from describing it to explaining it (Patton, 2012; Pawson & Tilley, 1997). To explain the real world, we must consider causality. Understanding how outcomes are caused in the real world is the central endeavour of a realist inquiry (Pawson & Tilley, 1997). This argument underpins Pawson and Tilley's (1997) realist evaluation, an approach to theory-based evaluation that is founded on the principles of realism.

While not known as a systems thinking methodology, realist evaluation considers programmes as open social systems. Pawson and Tilley (1997) recognised the complex interactions and relationships that influence how complex social processes and structures emerge over time. As Pawson and Tilley (1997, p. xiii) put it: "The reality we seek to explore is stratified".

This stratification is a result of social interaction, social structure, and personal agency, as well as how the interdependencies amongst them develop and emerge into real-world rules, institutions, norms, and customs. As such, the realist stance is that all human action is embedded within this expansive, layered range of social processes, rules, and institutions that emerge over time, and that this context informs how we interact with our reality (Pawson & Tilley, 1997), as shown in Figure 23.

Figure 23

Realist Context Surrounding the Programme



Note. Adapted from *Evidence-based Policy: A Realist Perspective* (p. 32), by R. Pawson, 2006, Sage Publications.

Copyright 2006 by R. Pawson.

Of interest to realist inquiries is the mechanisms that underlie causal explanation. Pawson and Tilley's (1997) realist approach is encapsulated by the following proposition: outcomes = mechanism + context, meaning that programme outcomes are caused by particular mechanisms that are activated by particular contexts.

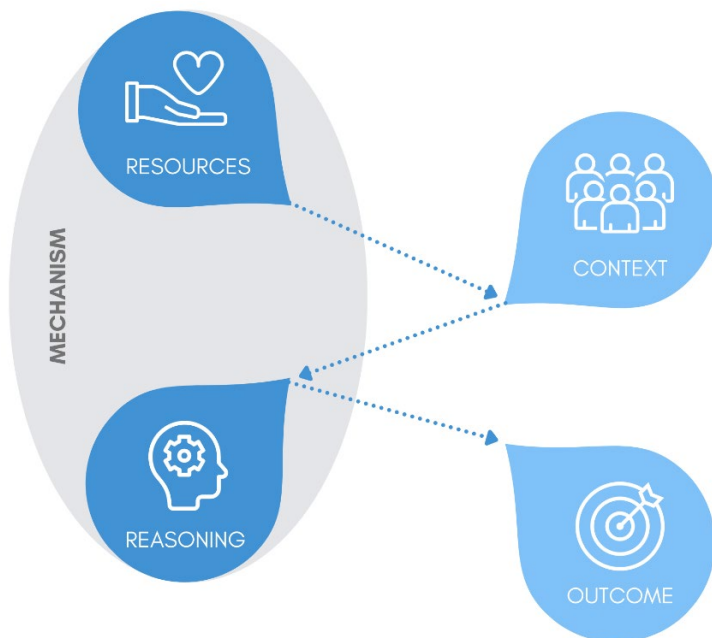
While outcomes are easily observable, mechanisms are less so; they represent the hidden inner workings and cognitive reasoning that determine how participants process and respond to a programme and the resources it offers.

Programmes, in this sense, do not work independently to produce change; rather, they work *through* participants by triggering internal transformation (mechanisms) in

particular contexts (see Figure 24). They provide participants with resources (e.g., knowledge, skills, physical resources, etc.) in a particular context, which then causes a change in the participants' reasoning, which, in turn, leads to an observable outcome (Dalkin et al., 2015). As such, programmes are opportunities for change that participants will either choose to take advantage of, or not, depending on the surrounding conditions and context (Pawson & Tilley, 1997).

Figure 24

Realist Mechanisms



Note. Adapted from “What’s in a Mechanism? Development of a Key Concept in Realist Evaluation”, by S. M. Dalkin, J. Greenhalgh, D. Jones, B. Cunningham, & M. Lhussier, 2015, *Implementation Science*, 10:49, 4 (DOI 10.1186/s13012-015-0237-x). Copyright 2015 by Dalkin et al.

Realist evaluation is method-neutral; it does not prescribe a particular set of steps. Rather, evaluators are encouraged to work it into their preferred theory-based evaluation practice. This involves hypothesising programme theory in the form of multiple context–

mechanism–outcome (CMO) patterns (using matrices), which are then tested and refined in the course of the evaluation. The idea is to answer the question ‘What works for whom in what circumstances, and why?’ CMO patterns can be developed in stakeholder workshops and by drawing on related literature and formal social science theories (Westhorp, 2014).

Realist evaluation adds value to ToC-building by prompting us to think critically and carefully about how outcomes are generated outside of direct programme influence. It can help to uncover blind spots in programme logic through purposeful consideration of potential mechanisms that are generally hidden from view, particularly the way in which participants’ reasoning is affected by the programme.

However, it is worth noting that realist evaluation is known to be a difficult approach, largely due to a lack of instruction and the conceptual challenge of identifying and separating mechanisms from context (Dalkin et al., 2015; Salter & Kothari, 2014).

Instead of embarking on a resource- and time-intensive foray into realist theory-building, the SAMEL approach draws on certain surface-level ideas to strengthen a ToC. This involves distinguishing between mechanisms and outcomes in the ToC, and highlighting the mechanisms that are context-bound. Brainstorming mechanisms during the development of the ToC could be augmented with social science and behavioural theory, i.e., pre-existing theories on the relationships between mechanisms and outcomes^x.

However, specifying the precise contextual conditions relevant to particular mechanisms may be difficult. This identification process could be enhanced by thoughtful

^x E.g., self-efficacy theory, theory of reasoned action, theory of planned behaviour, social cognitive theory, social learning theory, rational choice theory, innovation diffusion theory, the Health Belief Model, social action theory.

deliberation on the participants' specific context and the ways in which the layers of that context interact and inform the reality in which participants are embedded. The SAMEL Toolkit proposes the use of Kurt Lewin's (1936) life space model for this purpose (see also Burnes & Cooke, 2013).

Life Spaces. Gestalt psychologist Kurt Lewin developed field theory from the 1920s onwards as a way to understand individual behaviour and, later, as a method to analyse and change group behaviour (Burnes, 2007). Today, organisational development practitioners and change managers have adopted what Lewin scholars call a "watered down" (Burnes & Cooke, 2013, p. 409) version of Lewin's original field theory, commonly referred to as 'force field analysis'. Force field analysis identifies both the factors that drive action towards a given change and those that cause resistance to that change. Identifying and ranking these enabling and obstructing forces allows organisations to devise strategies to support the former and minimise the latter.

However, this interpretation of Lewin's field theory has come under fire for oversimplifying Lewin's original gestalt psychology principles (Burnes & Cooke, 2013; Swanson & Creed, 2014), which views human behaviour as resulting from how people perceive their environment, and how elements of that environment interact to further influence perception and behaviour.

Lewin believed that, to understand a change in behaviour, we must understand individuals' or groups' life space — the factors that make up their perceived environment or reality, together with the strength of the interactions between those factors (Lewin, 1936; see also Burnes & Cooke, 2013). Creating change or disrupting the status quo thus has to start with an understanding of the relevant forces, together with the interactions between them

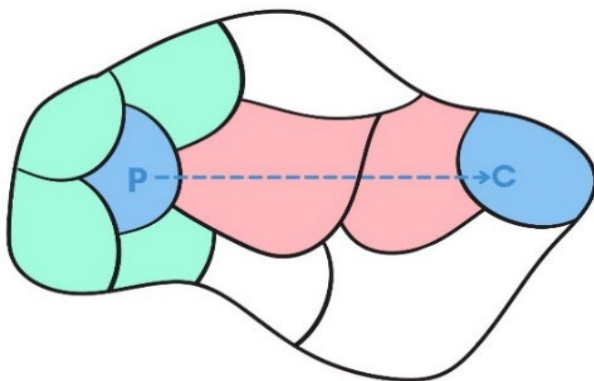
and the desired change. This is followed by identifying actions to increase supporting forces and decrease obstructive forces.

In developing a ToC, this becomes useful for understanding both micro- and macro-level contextual conditions that enable participant change. Programme teams may better understand contextual conditions by applying Lewin's (1936) life spaces to the realist relationship between the programme and participant change (seen in mechanisms and outcomes), i.e., how participants' context influences or constrains the way they learn and benefit from the programme.

An amended version of Lewin's skeleton for a life space is depicted in Figure 25, indicating the participant (*P*), the goal of participant change (*C*), and the journey from one to the other (the dotted line). It is important to note that, here, *participant* refers to those actively participating in the programme (as opposed to those who may ultimately benefit).

Figure 25

Life Space Model with Colour-coded Forces



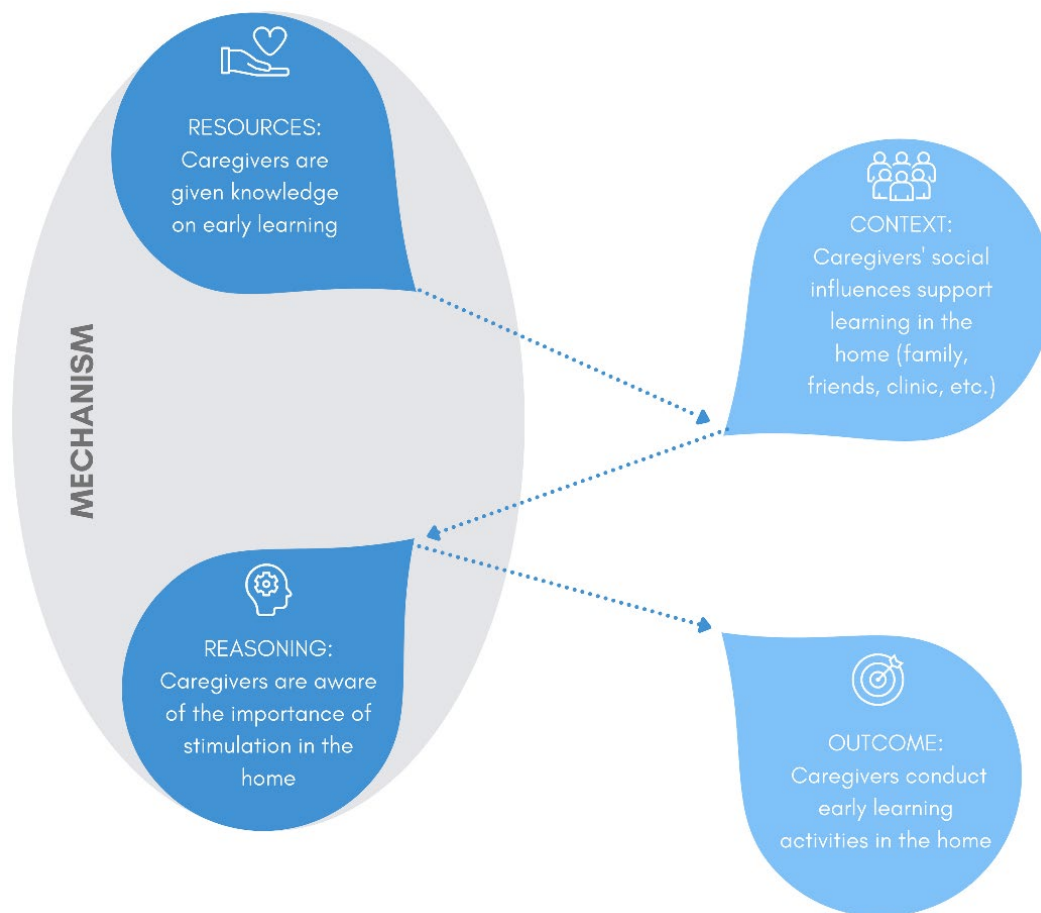
Note. Adapted from "Kurt Lewin's Field Theory: A Review and Re-evaluation", by B. Burnes & B. Cooke, 2013, *International Journal of Management Reviews* (15), p. 413 (DOI: 10.1111/j.1468-2370.2012.00348.x). Copyright 2012 by B. Burnes & B. Cooke.

The enabling factors that support participant change lie behind the participants, coloured green in the figure. The factors that surround the journey between the participant and change represent the forces that resist this change (coloured pink). Secondary forces (white) interact with and influence the enabling and resisting forces. The size of the sections in the life space represent the strength of a particular force. Given the interactive nature of the forces, as one grows in size, another may shrink. See Appendix Q for an example of an ECD-related life space model.

The SAMEL Toolkit provides a life space worksheet to practise mapping contextual conditions. To identify the life space conditions that should be inserted into a ToC, the Toolkit encourages evaluators to match contextual conditions to participant mechanisms, i.e., determining the context in which we could expect a participant's reasoning to change (see Figure 26 for an example). This is done by drafting CMO relationships in a practice worksheet.

Figure 26

Example of Matching Context to Mechanisms



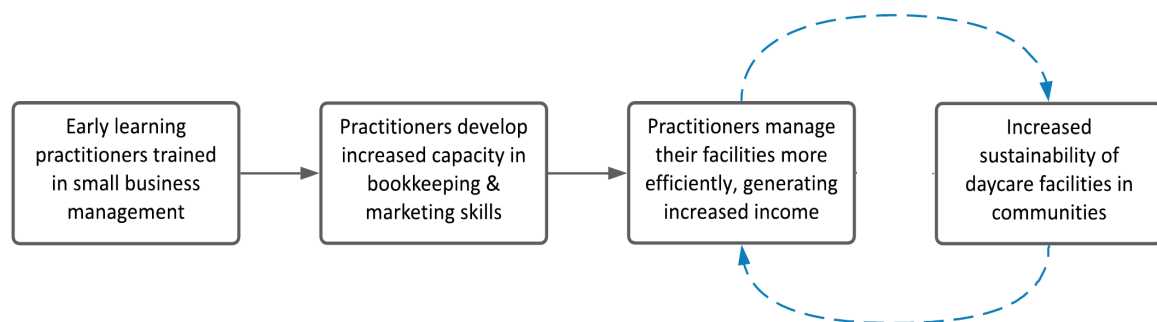
The relationships identified between the programme and its outcomes and mechanisms may be non-linear in nature. The Toolkit thus encourages evaluators to depict non-linear relationships in the ToC using feedback loops.

Feedback. Non-linear relationships representing the effect of feedback in a given system cannot be adequately captured by linear depictions of programme theory, such as matrices. Many authors have resorted to the other extreme, advocating completely circular illustrations, using CLDs (Dyehouse et al., 2009; Stroh, 2015). As described in the discussion of the Scoping chapter, CLDs are used in the field of system dynamics to illustrate the interplay between cause, effect, and feedback between variables in a given system.

System dynamics and the use of CLDs typically involve quantifying stock- and flow variables using computer modelling and mathematical equations in order to simulate manipulations of system behaviour, and thus model future behaviour, or to reveal patterns of behaviour (Forrester, 2007b). This is, of course, beyond the scope of ToC development. Instead, the SAMEL Toolkit encourages evaluators to highlight relevant feedback between ToC elements, such as between two outcomes. Reinforcing loops can be identified in situations where achieving one outcome has an amplifying effect on another outcome. Balancing loops can be identified in situations where a programme activity or outcome helps to reduce an undesired state or condition, which, in turn, helps to increase a particular outcome. Identifying feedback loops in a ToC can be especially beneficial in the identification of unintended outcomes. A simple example is provided in Figure 27.

Figure 27

Example of Reinforcing Feedback



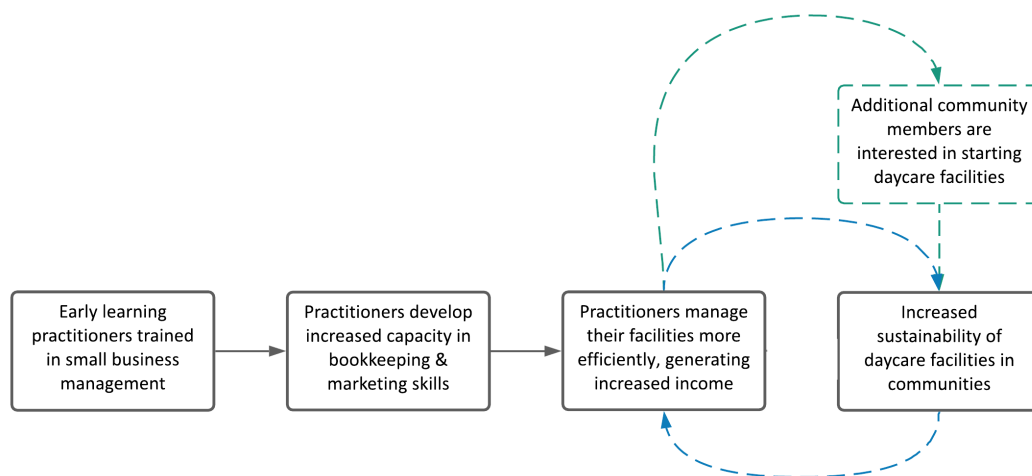
In this example, programme participants (early learning practitioners) experience greater capacity to manage the business side of their day care facilities, which then improves their ability to generate more income. This then increases the sustainability of communities' access to day care, and the facilities become a viable source of income. A reinforcement loop

(illustrated in blue) can be seen in this relationship; a higher income from the business means greater sustainability.

In this example, there is also the potential of an unintended outcome having both positive and negative effects on the sustainability of day care in a given community. This is illustrated in green in Figure 28.

Figure 28

Example of Unintended Feedback

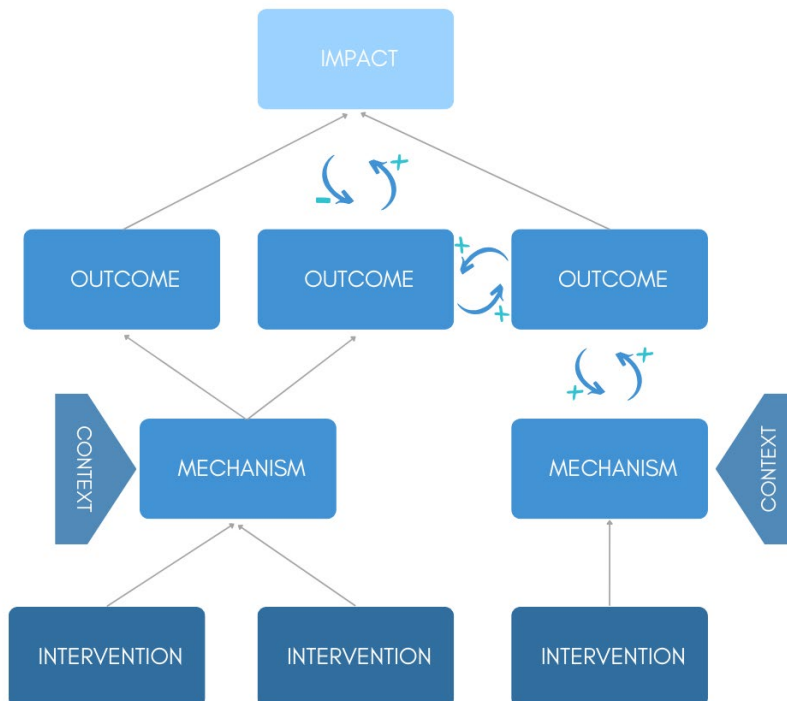


As programme participants generate greater income from their day care facilities, other community members may establish competing facilities. While this positively impacts the long-term objective of increasing the sustainability of access to day care in communities, competition may also lead to an oversupply of services, thereby reducing the income of practitioners.

While the SAMEL Toolkit encourages evaluators to depict important feedback relationships in ToCs (practised using a worksheet), it discourages over-using feedback loops and CLDs, as this may cause visual overload and confusion (Chapman et al., 2023; French et al., 2023). The Toolkit provides examples of linear depictions of a ToC that also incorporate non-linear relationships (see Figure 29).

Figure 29

Example of ToC Illustration using Non-linear Relationships



Theory of Action

The SAMEL ToA describes how the programme will go about activating the ToC and its stated change process (Chen, 2015; Funnel & Rogers, 2011), focusing on programme implementation. The ToA incorporates the following elements:

1. Participant definition: An operational definition of the specific group of people who will actively participate in the programme, including the criteria that make them eligible for the programme. If appropriate, a secondary target group of passive participants may also be defined (i.e., they do not actively participate, but passively benefit from what is achieved by the participants);
2. Outputs: The specific results or deliverables that are expected to result from the team's activities. Unlike outcomes, which are achieved by the programme

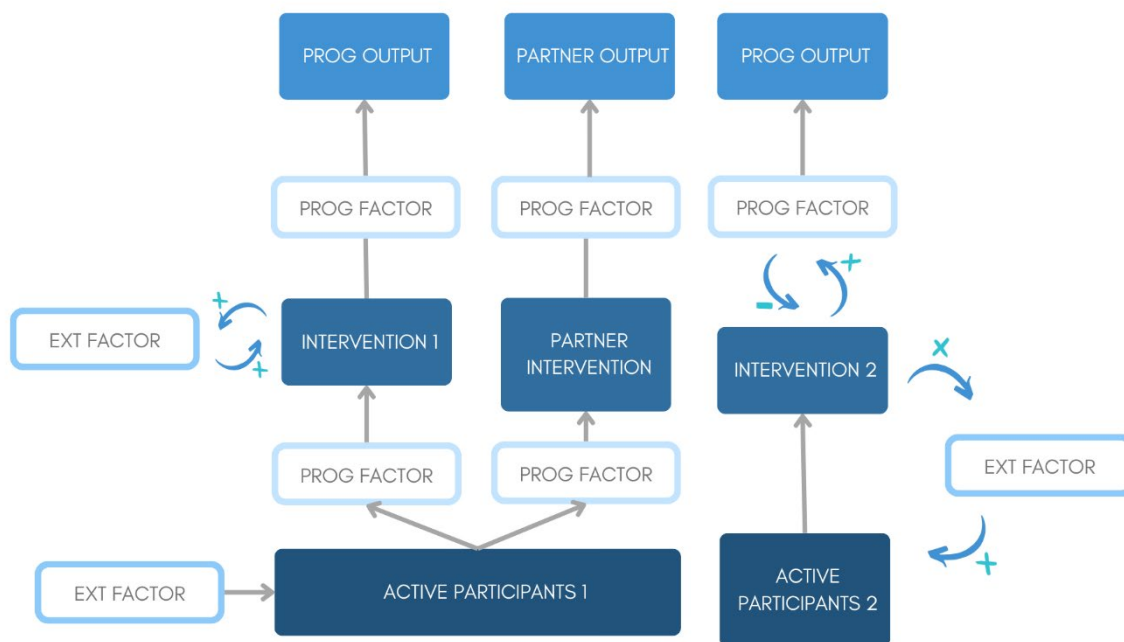
participants, outputs are achieved by the programme implementers and staff, and are needed for the interventions to take place;

3. Partners: Other actors with whom the programme collaborates (or intends to collaborate) to achieve particular outputs or contribute to participant outcomes;
4. External factors: Factors outside of the programme that may directly influence the programme’s implementation or the achievement of outcomes; and
5. Programme factors: Factors relating to the implementation of the programme that may directly influence the quality of implementation or participants’ engagement with the programme.

As with the ToC, evaluators are encouraged visually depict the ToA using a diagram (see Figure 30).

Figure 30

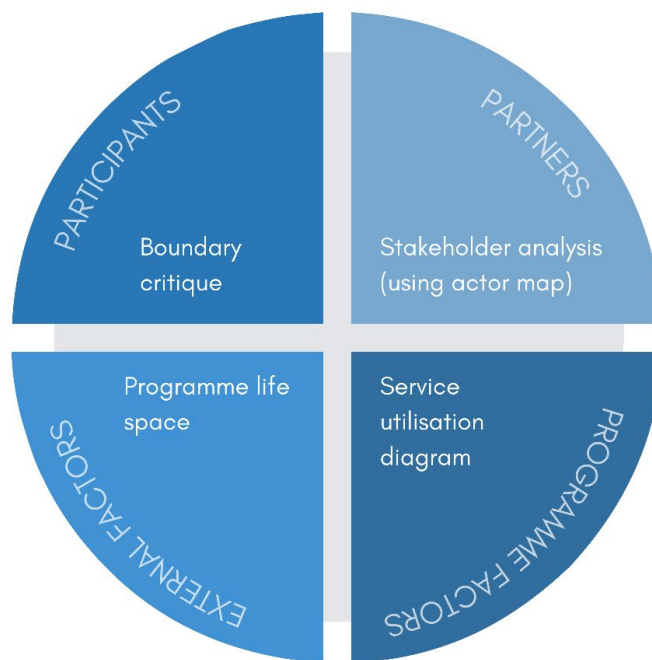
Example of a ToA Illustration



The identification of programme outputs is generally a straight-forward process in which the team's views and experiences are most valuable. To guide the identification of the rest of the ToA elements, the SAMEL Toolkit proposes four tools, which I sourced from the evaluation and systems thinking literature (see Figure 31), discussed below.

Figure 31

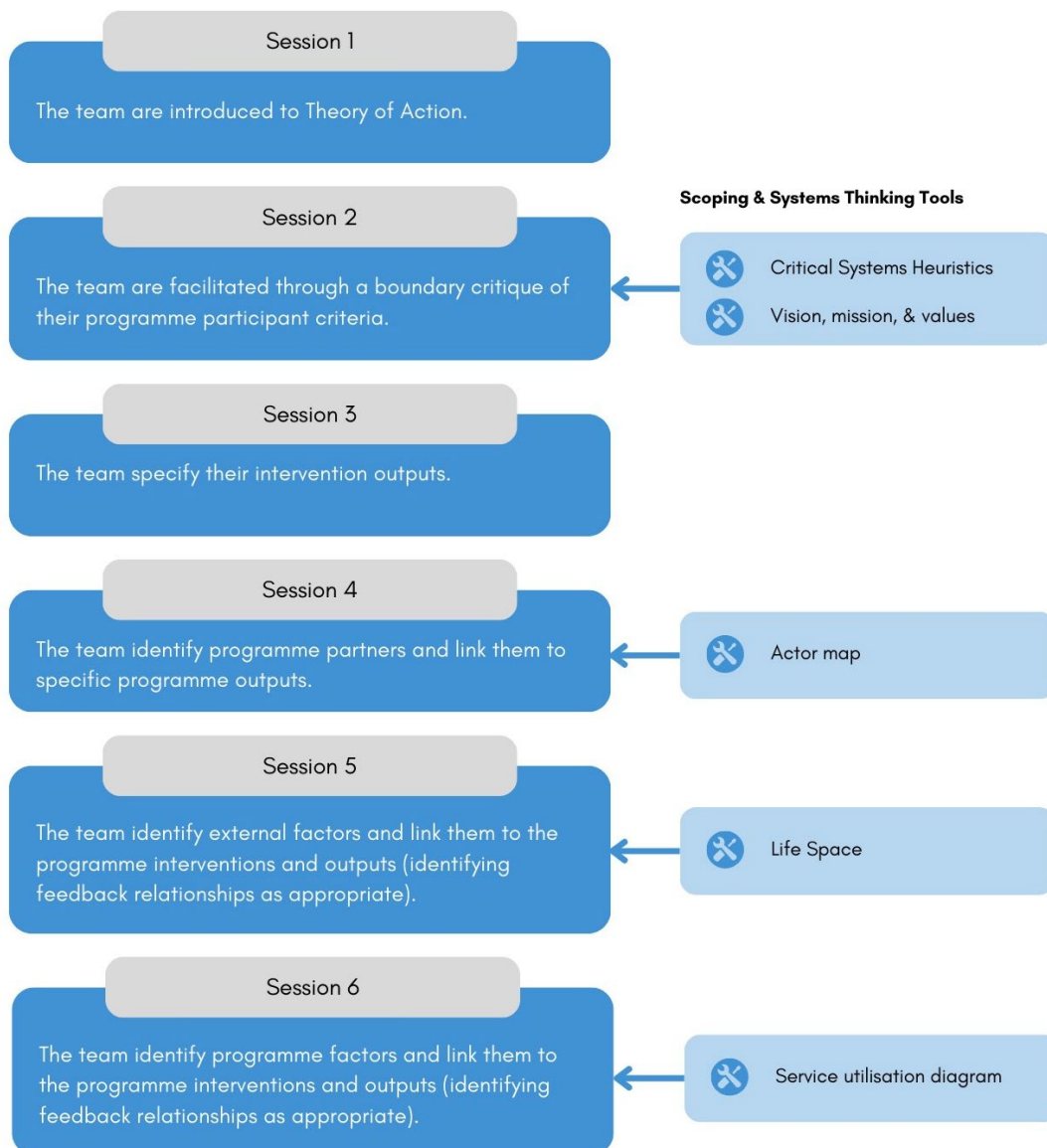
ToA Elements and Associated System Thinking Tools



The aim is to develop the ToA by formulating a definition of the participants, specifying outputs, linking partner contributions, and identifying both external and programme factors. The workshop sessions are summarised in Figure 32.

Figure 32

ToA Workshop: Summary of Sessions



In the following sections, I describe the systems thinking tools that the Toolkit recommends evaluators facilitate in Sessions 2, 4, 5, and 6 of the ToA workshop.

Boundary Critique (Critical Systems Heuristics). An important component of the programme’s focus is to firmly establish who will be served by the programme and, thus, who we can expect to undergo particular changes as a result. Drawing boundaries is inescapable; therefore, it is essential to be aware and critical of why and where boundaries are drawn

(Williams & Hummelbrunner, 2011). Doing so helps to test claims of inclusivity of social programmes and to foster critical awareness of the interests or voices that are given prominence (Ulrich & Reynolds, 2010).

Critical Systems Heuristics (CSH), a cornerstone of critical systems thinking, acts as a conceptual framework that challenges the idea of holism by noting the need for boundaries, and thus the need to critically deliberate on the decisions that inform boundaries (Williams & Hummelbrunner, 2011).

The central component of CSH is boundary critique — the systematic questioning of boundary choices for a particular system (Ulrich & Reynolds, 2010), such as a programme or policy. Questioning takes place in two modes: the (actual) *is* mode, which describes the system as it currently operates, and the (normative) *ought* mode, which describes how it should ideally operate. CSH provides a heuristic of 12 questions to guide this process, the answers to which describe how the system has been framed and bounded, illustrating the values, norms, and assumptions at play. CSH differentiates between those who have been involved in the system versus those who are excluded but are affected. This notion, together with normative versus descriptive questioning (*ought vs is*), is helpful in encouraging programme teams to think critically about who they ought to serve, as well as who they inadvertently marginalise through their boundary decisions. Ignoring these kinds of decisions can have ethical ramifications by potentially contributing to social inequities (Williams & Hummelbrunner, 2011).

While answering the CSH questions in the *is* mode is relatively straightforward, as a current situation is described, answering in the *ought* mode is more challenging. There are few guidelines for doing this. Some authors (e.g., Reynolds, 2006; Ulrich & Reynolds, 2010) suggest roughly answering *ought* questions yourself using reference material, background

information, personal familiarity with the situation, and your “initial impressions of the situation and associated system” (Reynolds, 2006, p. 111).

The SAMEL Toolkit advises evaluators to draw on the situation analysis and SSM root definitions to develop criteria for an ideal participant group or target population, including all relevant demographic variables that are applicable to the programme or problem of interest (e.g., region, age range, gender, education level, socio-economic status, and access to services).

This definition process should be aligned to the programme scope (problem statement, mission, and vision), with evaluators asking questions such as:

- Is there alignment?
- Does the participant group make sense in the context of the programme scope, and vice versa?

If the programme is established, the *ought* can be compared with the *is*:

- What is the difference between the ideal and the current participant group?
- What is the practical consequence of this difference?
- What does the difference indicate about the programme team’s worldview and assumptions?

Lastly, the boundary choices can be considered in light of the programme values that were agreed upon for the programme scope.

- Are the boundary choices made by the team in line with the programme values?

Once the participant criteria have been determined, the Toolkit advises reflection on those groups who fall outside of those criteria but are still directly or indirectly affected by the problem. Discussion and debate should centre on the reasons for their exclusion, as well as

justification for these reasons. The findings of the situation analysis can be used in this exercise. For example, the root definitions may reveal that the marginalised in this case are the participants (or customers) of a different programme. The Toolkit provides a worksheet for evaluators to prepare for this exercise by reflecting on the situation analysis and participant characteristics.

Revisiting Actor Maps. In the event that a programme needs particular community and associate partners to implement interventions or access services that are important in achieving the desired participant outcomes, the ToA should identify who these partners are and what strategies are needed to establish these relationships (Chen, 2015). The Toolkit advises evaluators to help programme teams reflect on potential programme partners by drawing on the actor map created in the Scoping workshop. A worksheet is provided to help teams identify five types of programme partners (Chen, 2015; Ulrich & Reynolds, 2010):

1. Core organisations that the programme team needs to work with closely to successfully implement the programme;
2. Related organisations that have the power to interfere with the programme's implementation, either purposefully or inadvertently;
3. Auxiliary service organisations that offer important additional services that programme participants may need;
4. Collaborating organisations that deliver similar or related services, and may offer implementation support through co-ordination; and
5. Marginalised organisations that represent the voices of those who have been excluded from the programme but are still affected by it.

Life Spaces. The Toolkit proposes again drawing on life spaces (Lewin, 1936) to assist programme teams in identifying external factors. External factors are those contextual

aspects that lie outside of the programme and its operations but are powerful enough to influence the programme's implementation. This encompasses influence at the macro-level, referring to the community's interest, support, or buy-in (Chen, 2015). Drawing up life spaces for a programme's macro-level support is intended to help the programme team reflect on enabling and obstructive forces that need to be attended to, as well as the interactions between these forces and others in the environment as a whole. Prior to the ToA workshop, evaluators are advised to draw an example/draft model based on their knowledge of the programme and the situation analysis, using a worksheet.

To incorporate the external factors into the ToA, the Toolkit advises identifying the most important factors that might influence or constrain the programme's implementation at the participant level or intervention level:

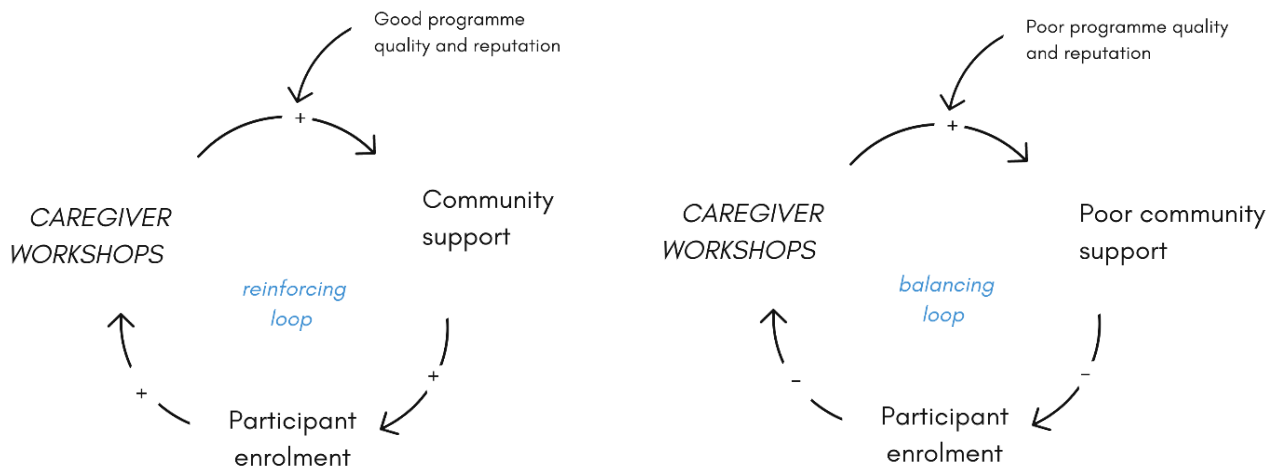
- Participants: What external factors might influence whether the participants access or participate in the programme (i.e., issues of recruitment, access, and attendance)?; and
- Interventions: What external factors might influence the interventions taking place (i.e., issues of service delivery and quality of delivery)?

Evaluators are also advised to consider feedback relationships between these factors and ToA elements, and illustrate these relationships in the ToA (see Figure 30). For example, community support for the programme might be identified as a critical external factor for participant entry into the programme. If the community is aware and supportive of the programme, the targeted participants will be more likely to enrol. This is illustrated in Figure 33. There may be a reinforcing relationship between this factor and the programme: the longer the programme is delivered, the greater the community support may become, thus

increasing the pool of interested participants. However, the quality and reputation of the programme plays a role in this relationship.

Figure 33

Reinforcing vs Balancing Effect of Programme Quality and Reputation on Participant Enrolment



Service Utilisation Diagrams. The external factors identified previously constitute the macro-level conditions that may support or inhibit implementation, whereas programme factors are associated with programme delivery, and are largely within the control of the programme team. However, these factors do have the power to disrupt implementation goals and, thus, the realisation of outcomes.

According to system dynamics, effective systems leverage effective stock- and flow mechanisms — the structural design of how systems accumulate and move ‘stocks’ (units of interest) through the system (Meadows, 2009). This concept can be applied to programmes, where the 'stocks' are the participants, and the 'flows' represent their progression through the programme. A service utilisation flowchart, as discussed by Rossi et al. (2019), serves as a visual tool to map out this participant journey, highlighting critical stages of participation and potential points of attrition. By depicting the entire pathway from initial contact to

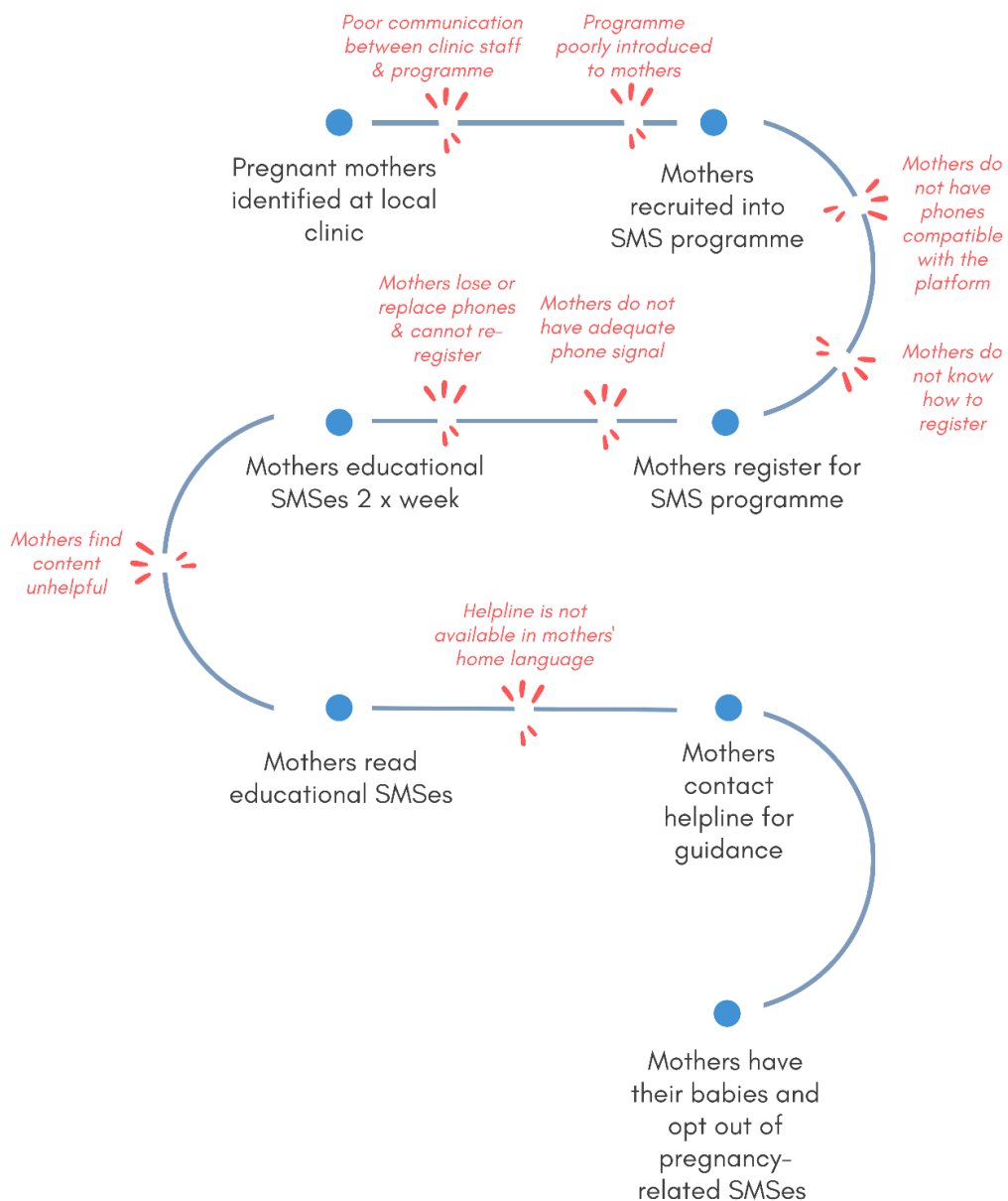
programme completion, the flowchart communicates the principles of stock and flow by tracking how participants ('stocks') move through different stages ('flows') of the programme. Thus, both frameworks aim to optimise system efficiency and effectiveness by identifying and addressing bottlenecks and barriers that impede smooth progression.

The Toolkit proposes the use of a service utilisation diagram, inspired by Rossi et al.'s (2019) flowchart, to visually illustrate the assumptions about how and why programme participants become engaged with the programme in the first place, and how engagement is maintained until it is intended to terminate (Rossi et al., 2019). The diagram plots the participants' programme journey from the very first point of contact to their graduation from the programme (Rossi et al., 2019). The flowchart includes possible points of departure from programme engagement, where participants may encounter bias or barriers to programme access. The counterpoint to these obstructions become the programme factors that are incorporated into the ToA.

Figure 34 is an example provided in the Toolkit, which was inspired by the real-life South African programme *MomConnect* (LeFevre et al., 2018; Peter et al., 2018). *MomConnect* is an SMS service with which pregnant mothers register during clinic visits. The service delivers educational SMSs to mothers throughout their pregnancy. The programme's evaluation findings (LeFevre et al., 2018; Peter et al., 2018) were drawn upon to help populate the diagram overleaf.

Figure 34

Service Utilisation Diagram of the MomConnect Programme



Based on the diagram, a number of programme factors can be identified that are critical in ensuring that mothers are supported to move through the programme pipeline as intended. The factors largely within control of the programme are: good working relationships with clinics; clinic staff are equipped to adequately introduce the programme to mothers; the programme is compatible with all types of mobile phones; the registration process is easy; the registration process allows for profiles to be connected to new devices; SMS content is

aligned with the needs of mothers and is presented in an engaging manner; and mothers can access helpline support in their home language.

Using service utilisation diagrams, the identification of programme factors is kept streamlined and relevant through a focus on priority factors that (a) are potential threats to programme success and (b) can be influenced through strategic decision-making. Evaluators are provided with a template to plot the flow of participants' service utilisation.

Expert Feedback on Focusing Chapter

As with the Scoping chapter, this chapter was well received by the three experts. Sarah, the systems thinking expert, liked the incorporation of realist evaluation, system archetypes, and feedback loops into the programme theory. Sarah noted that the Scoping and Focusing chapters worked together well, saying: *“For non-profit staff, this seems incredibly useful to apply systems thinking to design.”* Naledi and Allette also particularly liked these two chapters, with Naledi commenting that they were *“strong on programme design skills for evaluators”*, which Naledi noted was an important skill.

A notable critique from Sarah was that the Focusing chapter relied on traditional programme logic and did not require that the team completely revise the intervention strategy. She noted that, rather, the content seemed to help the team improve what is already being done.

“The maintenance of programme logic and process constrains the potential value for coming up with totally new ways of framing problems/situations, new goals, new strategies/activities, etc. Rather, the main value seems to be expanding peoples' views on what they already do to be more responsive to the situation, inclusive of multiple perspectives, purposeful about their change process, and aware of the importance of MEL.” (Sarah)

I agree with this critique. However, while the Scoping and Focusing chapters are heavily geared towards programmatic design, the SAMEL Toolkit is ultimately a MEL product. As such, I had to make significant trade-offs in terms of focus and time afforded to particular sections. To keep the focus on MEL, it was deemed impossible to support complete programmatic revision in the workshop, as Sarah suggested. To do so would be valuable and maximally beneficial in terms of what systems thinking can offer programme design, but this was a pragmatic trade-off that needed to be made to keep the workshops manageable and primarily focused on MEL.

Alletta and Naledi had specific suggestions for changes, of which the following were incorporated:

1. Worksheet instructions were expanded and clarified.
2. CSH and feedback worksheets were created.
3. Stakeholder analyses were streamlined between workshops, to avoid repetition.
4. Programme participants (ToA) were split into *active* and *passive* to account for the different levels of populations that benefit from programmes.
5. Workshop guidelines were amended to combine the sessions that develop mechanisms and outcomes.

Toolkit Chapter 4: Monitoring & Learning

The key objectives of the Monitoring & Learning chapter are to guide evaluators in the development of the NPO's programme monitoring plan and learning strategy. This is done by reflecting on the data needed for adequate management and ongoing improvement of the programme, and then deciding how this data will be collected, used, and fed back into the programme's management.

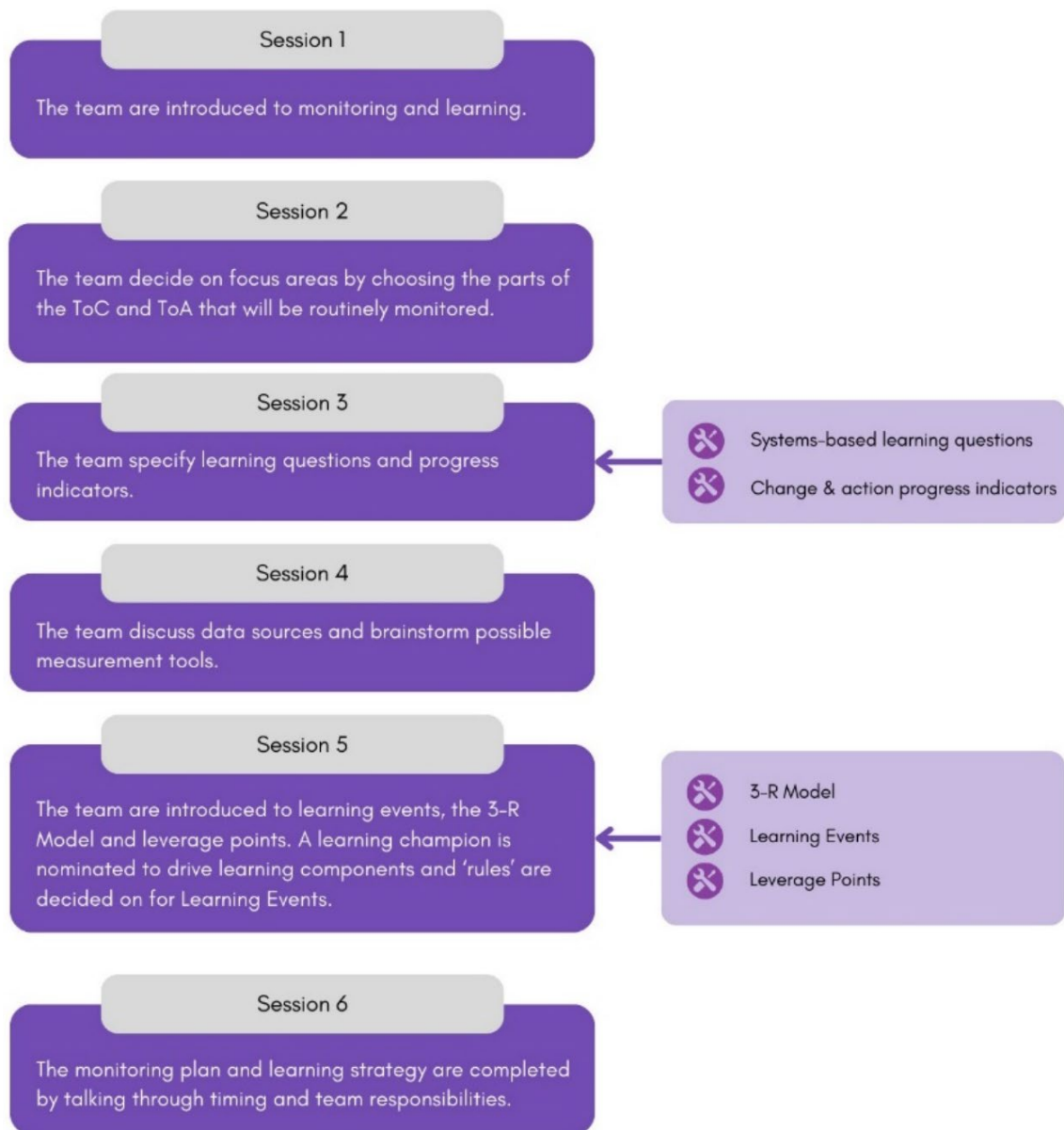
Evaluators are advised to collect information from team members on their data needs and available resources for monitoring prior to the Monitoring & Learning workshop.

Evaluators are also advised to draft indicators, data sources, and measurement tools so that workshop discussions can efficiently blend the team's insights and knowledge with the evaluator's technical expertise.

The Monitoring & Learning workshop is also afforded a full day, and uses the programme theory as its guiding resource. The sessions of the workshop are provided in Figure 35.

Figure 35

Monitoring & Learning Workshop: Summary of Sessions



Monitoring

The programme monitoring system uses the programme theory as a foundation. The ToC guides the development of the outcoming monitoring component, while the ToA guides the implementation monitoring component. The key task of the monitoring system is to collect data relevant to the programme theory to test the theory against reality.

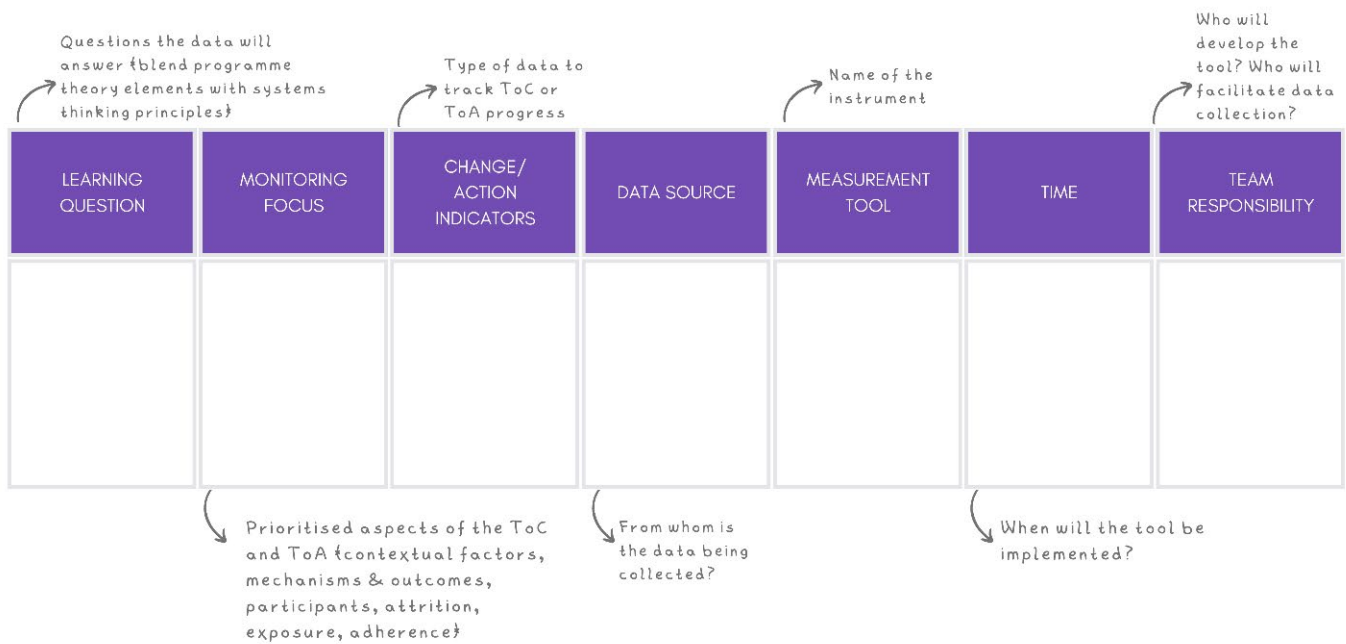
The SAMEL Toolkit differentiates monitoring from evaluation with regard to purpose. Monitoring is used to track programme implementation and participant outcomes, while evaluation is proposed as a tool for delving deeper into issues identified through monitoring, or for topics logistically suited to monitoring (such as assessing the outcomes of passive participants or wider community benefits).

The monitoring aspect of the Toolkit is traditional in the sense that it primarily draws on the work done in the Focusing workshops to develop a monitoring plan (see Figure 36). The chapter provides guidance on developing a monitoring plan, including practical advice and considerations regarding measurement tools and data collection.

Given that the monitoring plan proposed is relatively traditional, the Monitoring & Learning chapter does not explicitly use systems thinking tools, but rather focuses on measuring the systems-based programme theory. The chapter does, however, introduce three additional concepts to broaden evaluators' and teams' thinking around monitoring. These are shown in Figure 36 and described below.

Figure 36

Toolkit Excerpt: SAMEL Monitoring Plan Summary



Systems-based Learning Questions. Learning questions help guide the collection and analysis of data deemed relevant to the programme (Markiewicz & Patrick, 2016). The SAMEL Toolkit proposes developing learning questions using the systems concepts. This is intended to broaden traditional MEL data collection to include systemic components, such as the programme’s relationships with partners and participants’ perceptions of the programme (see Figure 37).

Figure 37

Toolkit Excerpt: Learning Questions using Systems Concepts

	Theory of Change	Theory of Action
<i>Boundaries</i>	<p>What effect are participants' contexts having on their ability to benefit from the programme? (How permeable is the programme-context boundary?)</p> <p><i>Focus area: Contextual factors</i></p>	<p>What effect is the programme's context having on its implementation?</p> <p><i>Focus area: External factors</i></p> <p>To what extent is the programme enacting its participant boundaries?</p> <p><i>Focus areas: Participants (involved and disengaged)</i></p>
<i>Systems and parts</i>	<p>To what extent is the programme system achieving its change-related purpose (mechanisms and outcomes)?</p> <p><i>Focus areas: Mechanisms and outcomes</i></p>	<p>How effectively is the programme system working to deliver its intended objectives?</p> <p>What is enabling or impeding programme implementation?</p> <p><i>Focus areas: Outputs, exposure, adherence, programme factors</i></p>
<i>Relationships</i>	<p>What causal relationships are being observed among participants and the programme?</p> <p><i>Focus areas: Mechanisms, outcomes, feedback</i></p>	<p>What causal relationships are being observed between the programme and other factors?</p> <p><i>Focus areas: All, feedback</i></p> <p>How effective are the programme's existing partnerships?</p> <p><i>Focus area: Partners</i></p>
<i>Perspectives</i>	<p>How do participants perceive the benefits of the programme, if any?</p> <p><i>Focus areas: Mechanisms, outcomes (+unintended)</i></p>	<p>How do participants perceive the programme's service delivery?</p> <p><i>Focus area: Participants</i></p>

Progress Indicators. The SAMEL Toolkit differentiates between action indicators, which are designed to track ToA components, and change indicators, which are designed to track ToC components.

In the tradition of realist evaluation, change indicators target the ToC's mechanisms, outcomes, and context in order to understand whether the programme is working (by achieving mechanisms and outcomes), and in what (contextual) conditions. The Toolkit

provides examples of indicators, as well as the types of data that could be collected. A resource is also provided that summarises a number of standardised, validated ECD tools that have been used successfully in South African contexts (including examples of their use).

A type of change indicator that may be necessary for programmes that intervene at a macro level (e.g., advocacy or policy change) are called ‘sentinel indicators’, which act as an alert for change taking place in the wider system. Sentinel indicators are proxies as opposed to direct indicators. They indicate whether a particular change appears to be taking place that warrants further investigation or measurement, and are appropriate to use when direct observation is not possible in the foreseeable future.

Many evaluators make use of SMART (specific, measurable, achievable, relevant, time-bound) or CREAM (clear, relevant, economic, adequate, monitorable) criteria to guide the development of indicators. The SAMEL Toolkit proposes a different approach — CRISP:

- Clear: Indicators should be concise and easily understandable to all team members.
- Relevant: Indicators should directly reflect ToC and ToA elements, and be relevant to *actionable* decision-making.
- Inclusive: Both quantitative and qualitative indicators should be used, to encourage well-rounded monitoring of focus areas, and to obtain in-depth data that can inform targeted decision-making.
- Sufficient (the ‘good enough’ principle): No indicator can capture the complexity of social change or the efforts of the programme team. Instead, indicators should be sufficient to alert the programme team to implementation issues that need correction.
- Pragmatic: Indicators should be feasible, realistic, and purpose-driven in light of the resources available for monitoring (staff capacity, time, and financial resources).

Participant Perspectives. While it is important to target the mechanisms and outcomes described in the ToC during outcome monitoring, it is also critical to be open to unanticipated and emergent changes that may occur amongst the participants, both positive and negative, as reported by the participants. This ensures that the programme team is not narrowly focused on particular predetermined effects and overlooking unanticipated changes. The Toolkit encourages evaluators to include open-ended questions, directed at participants, to probe how the programme has negatively or positively impacted them.

The 'In' and 'Out'. Monitoring who is participating in the programme is a component of ToA monitoring. This helps the team to establish whether the intended participants of the programme are being correctly targeted and appropriately involved in the programme. In addition to this, the Toolkit proposes monitoring those who have disengaged from the programme.

As discussed, the Focusing chapter of the Toolkit introduces CSH and the notion of boundary critique (Ulrich 1983). CSH features in the Monitoring & Learning chapter as well, guiding programme teams to collect data on the people directly involved and participating in the programme, i.e., demographic and enrolment data that indicate the effectiveness of participant recruitment. Additionally, CSH principles prompt teams to consider those who are indirectly affected or marginalised. In the context of actual programme implementation, the latter can refer to three things: (1) those who were initially enrolled in the programme and then dropped out (the disengaged), (2) those who fit the programme's selection criteria but were not selected, or (3) those who did not fit the selection criteria but are affected by the programme in some way. To keep this section manageable and actionable, the Toolkit recommends focusing on the first group, the disengaged. The latter groups may be best served by a more in-depth evaluation.

Programme attrition is a normal and expected outcome of implementation.

Participants may relocate, lose interest, or decide that the programme is not suitable for them. In other situations, interested participants may be inadvertently barred from participating in the programme by some obstacle (such as those theorised in the ToA's service utilisation diagram). This is an actionable issue that speaks to the programme's ethical obligation to the target population. It is thus important to understand *why* participants are dropping out, and whether the reasons for doing so are related to the programme's structure.

Learning

Integrating monitoring data into programme planning and management is arguably the most important component of monitoring efforts. Data collection and reporting are resource-intensive exercises that are justifiable only insofar as they are put to instrumental use.

It is therefore recommended that a learning strategy (see Figure 38) be developed as part of the programme's monitoring and learning framework, with the aim of formally mapping out how monitoring data will be integrated into programme management. The Toolkit suggests that the learning strategy include:

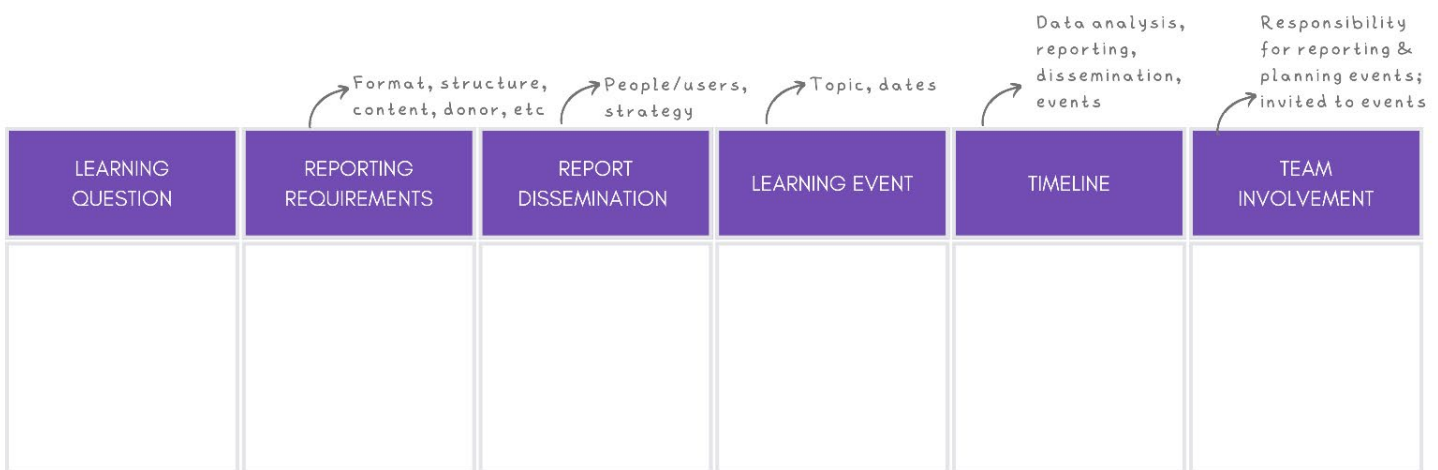
1. Reporting requirements: How will the monitoring data be reported? What form will reports take? How will the reports be structured? What reports do donors require?;
2. Report dissemination: With whom should reports be shared? How can data be fed back into the programme? How can data be communicated to a wider audience?;
3. Learning events: What will the topics of learning events be? How often will they take place?; and

Team involvement: Who will manage programme data? Who will compile monitoring reports? Who will attend learning events?

In addition to the development of a learning strategy, the workshop sessions are intended to provide the programme team with learning tools and concepts that they can implement throughout the programme year/monitoring cycle.

Figure 38

Toolkit Excerpt: SAMEL Learning Strategy Summary

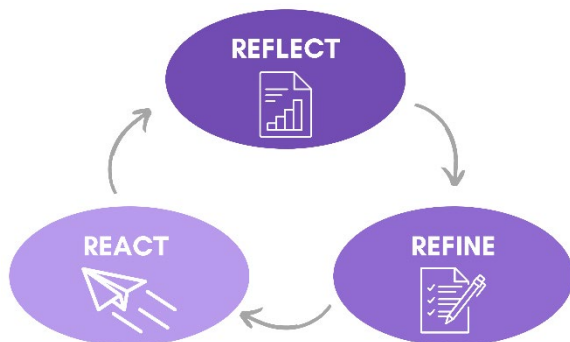


To practically support the integration of learning into the NPO’s MEL system, the Toolkit provides two learning mechanisms: learning events and the 3R Model, which I formulated for the Toolkit. In the Monitoring & Learning workshop, the team is introduced to the tools and capacitated to implement them throughout the year.

Learning Events. Learning events involve the regular coming together of programme staff to reflect on monitoring data and other important information, prioritise areas of interest, and agree on any corrective actions. To ensure that efforts are co-ordinated and structured, the learning events could be organised according to the 3R Model: reflect, refine, and react (Figure 39).

Figure 39

The 3R Model



In the first phase of the event, the team reflects on the monitoring data, paying attention to emergent trends and areas of interest. The focus in this phase is on what is working and not working with regard to implementation efforts and participant change. It is also important to reflect on the programme theory, and determine if and where it is falling short. Examples of reflective questions are:

- What is working?
- What is not working?
- What are we doing well?
- Where are we falling short?
- What has unexpectedly emerged?
- Which parts of our programme theory are being challenged, and why?
- Is our data collection strategy performing well? Does it need revision?

In the second phase of the event, the team refines these areas of interest to actionable priority areas. In this phase, the emphasis is on the most salient issues that need attending to in order to improve or sustain service delivery and participant benefits.

The Toolkit recommends that the list of priorities be kept to a maximum of five. In terms of social-ecological systems research, this is known as the 'rule of hand' (Holling & Sundstrom, 2015; Walker et al., 2006). This proposition states that we generally need no more than five variables to explain or understand the patterns and dynamics of system behaviour. Therefore, the programme team is challenged to conceptually distil their findings and reflections to a small set of issues that they think best represent the critical changes and patterns they have observed.

Guiding questions here are: (1) What five issues best represent our findings; and reflections? (2) What five issues can we feasibly attend to in the short term?

Lastly, in the third phase of the event, the team reacts to the data by putting together action plans for each priority — key tasks and deliverables, and for the parties responsible for them. The Toolkit provides a reflective worksheet, which incorporates the 3R Model, that evaluators can share with programme teams to use in future learning events.

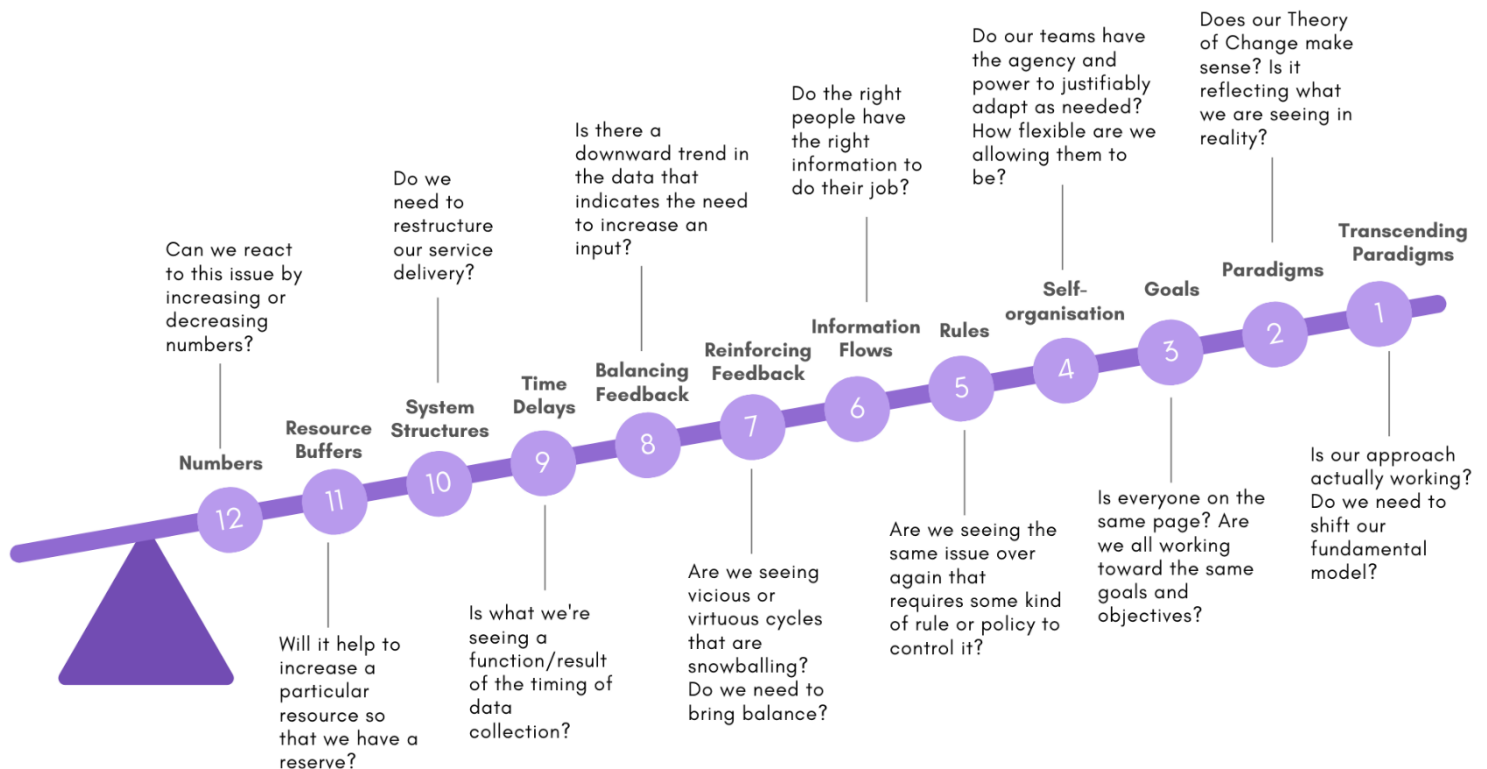
Reacting using Leverage Points. In order for the team to formulate action plans in future learning events (react phase), the Toolkit proposes that teams be capacitated to use leverage points as a heuristic for identifying strategic priorities. Leverage points are small, targeted actions that can produce widespread change in system behaviour (Meadows, 2009; Senge, 2006).

Seminal systems thinking author Meadows (2009) described 12 generic leverage points that can be found in any given system (such as an organisation or policy system). These are provided in the Toolkit so that evaluators can assist teams in considering particular areas in their programme and implementation strategies that might be responsible for the trends and patterns they are observing in the monitoring data. As such, Meadows's (2009) leverage points act as sensitising concepts that help structure the team's reflections on data in learning

events. Meadows (2009) lists the leverage points in order of power. This is illustrated in Figure 40, along with the reflective questions provided to the teams in a Toolkit worksheet, which evaluators can give to teams to use in future learning events.

Figure 40

Meadows's Leverage Points Illustrated with SAMEL Reflective Questions



Types of Learning Events. The 3R Model can be implemented in different types of learning events. The Toolkit offers three types:

1. General monitoring: The team reflects on recent implementation and outcome monitoring data.
2. Force field events: The team reflects on external and programme factors, sharing insights on how the context is influencing the programme. The Toolkit suggests that these events centre around the ToA's service utilisation diagram and life space model, with team members sharing insights from the field (i.e., how these models fare in real

life). Force field events include the effectiveness of programme partnerships, and a SWOT (strengths, weaknesses, opportunities, and threats) analysis is suggested to address this aspect (the Toolkit includes a SWOT worksheet).

3. Sectoral events: The team shares learnings from recent happenings in the sector that have a bearing on the programme. Programme staff should make an effort to participate in sectoral forums and communities of practice to stay informed about important events and changes that are taking place within the sector, and to learn about how these may affect the programme. The Toolkit encourages these learning behaviours, and suggests that team members share their learnings during these learning events.

Expert Feedback on Monitoring & Learning Chapter

Sarah deemed this chapter and the Evaluation chapter the weakest, with Alletta and Naledi offering the most suggestions for improvement. Nevertheless, Alletta felt it was a particularly useful tool for developing learning competencies amongst NPO teams. Likewise, Naledi appreciated the inclusion of the learning section in the Toolkit, saying that *“organisations still struggle with this”*.

The Toolkit originally did not discuss practical issues of data collection methods, which Sarah and Alletta both considered a concerning gap. This was addressed in the revised version of the Toolkit by including additional practical considerations for data collection (such as South Africa’s Protection of Personal Information Act 4 of 2013 and issues around child measurement). Additional suggestions from Alletta and Naledi that were incorporated are the following:

1. Guidance was incorporated on reliability and validity of data collection tools;

2. Clarity was improved around outcome monitoring (particularly in terms of monitoring context);
3. Learning events were more deliberately linked to the programme theory (e.g., reflection on service utilisation); and
4. Reporting was included in the learning strategy.

Toolkit Chapter 5: Evaluation

The key objectives of the Evaluation chapter are to develop an evaluation design, introduce key boundary choices in preparation for an evaluation, and draft an ideal evaluation plan that can later be fed into terms of reference. Introducing these concepts to the programme team during the development of the MEL framework helps to ensure they are prepared for the requirements of a future evaluation and feel confident to take on such a process. The ultimate objective is to reflect on the team's readiness to conduct an evaluation. Thus, the central goal of the Evaluation workshop is to adequately demonstrate to the programme team what a future in-depth evaluation will look like, and to help the team prepare and feel ready for it. The workshop sessions are shown in Figure 41.

Figure 41

SAMEL Evaluation Workshop: Summary of Sessions



As with all the Toolkit chapters, evaluators are advised regarding the preparation needed before the workshop (e.g., brainstorming criteria definitions and evaluation methods, preparing board game materials, etc.).

The SAMEL Toolkit employs hybrid evaluation, in which internal expertise and insights are combined with the scrutiny and expertise of an external evaluation consultant. This approach was chosen based on the following assumptions: (1) An external actor's assessment of the programme's value will likely, in the eyes of other stakeholders, have more legitimacy and power than that of the internal programme team, and (2) an external evaluator can work with

the programme team in a participatory manner to produce a high-quality evaluation (e.g., Bourgeois et al., 2011).

The Toolkit introduces evaluators to MAE. Chilisa and Mertens's (2021) nine MAE principles are woven throughout the chapter to ensure that evaluators incorporate them in discussions amongst the programme team in the workshop, and that the team is able to apply them in conducting future evaluations.

In the following sections, I discuss the evaluation components that are developed during the Evaluation workshop, together with the concepts that I incorporated into the Toolkit to aid their development.

Evaluation Design

To complete the first objective of the Evaluation workshop, the Toolkit provides guidance on evaluation design, encompassing evaluation purpose, criteria, questions, methods, and boundaries. The Toolkit provides two major systems-based tools: SAMEL evaluation criteria (which I formulated for the Toolkit) and the use of CSH in boundary-setting.

Evaluation Criteria. Should the purpose of an evaluation be summative, the Toolkit proposes the traditional use of criteria to enable normative judgments of the programme's value. Pre-existing evaluation criteria are often used to frame areas that are to be investigated in an evaluation (Markiewicz & Patrick, 2016), such as those promoted by international development agencies like the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) (2021). The DAC criteria are used in evaluation to determine the merit or worth of a programme, with each criterion representing a particular lens through which this judgment is made (OECD, 2019). The criteria play a normative role, describing the ideal or desired characteristics of programmes. The DAC proposes that the criteria be used thoughtfully, to support useful and high-quality

evaluations, and recommends that the use of specific criteria depend on the evaluation's purpose (OECD, 2021).

The DAC (OECD, 2021) criteria are useful and well-established starting points for the design of an evaluation. However, the evaluation questions that criteria inspire often fail to adequately capture the systemic concepts and the MAE principles. Furthermore, there is much debate around the usefulness and relevance of the DAC criteria in the post-COVID world, which is in dire need of social and environmental transformation (Patton, 2021).

While the shortcomings of the DAC (OECD, 2021) criteria may appear to simply suggest that tailored and context-specific criteria be developed for particular evaluations, it is important to have some degree of uniformity across programmes and evaluations, to enable intra-sectoral comparison and cross-organisational learning (Faust & Verspohl, 2018; Ofir, 2017b). Comparing programmes along the same or similar criteria enables the generation of a national database that promotes collective learning and system growth within the sector.

Alternative criteria to those of the DAC (OECD, 2021) have been proposed by evaluation experts, most notably Ofir (2017a, 2017c) and Patton (2021).

While Patton's (2021) criteria focus on global transformation, Ofir's (2017a, 2017c) focus on evaluation for development, and highlights perspectives from the Global South and complex systems. Ofir offers three levels of criteria: non-negotiable (criteria that should be included in all evaluations), flexible (criteria that could be included, depending on context and circumstances), and optional (criteria whose inclusion depends on stakeholders' particular interests at a given point in time).

The SAMEL Toolkit proposes a combination of all three approaches: DAC's (OECD, 2021) uniform and global perspective, Patton's (2021) transformational perspective, and Ofir's (2017a, 2017c) focus on complexity and the African context. The SAMEL criteria are:

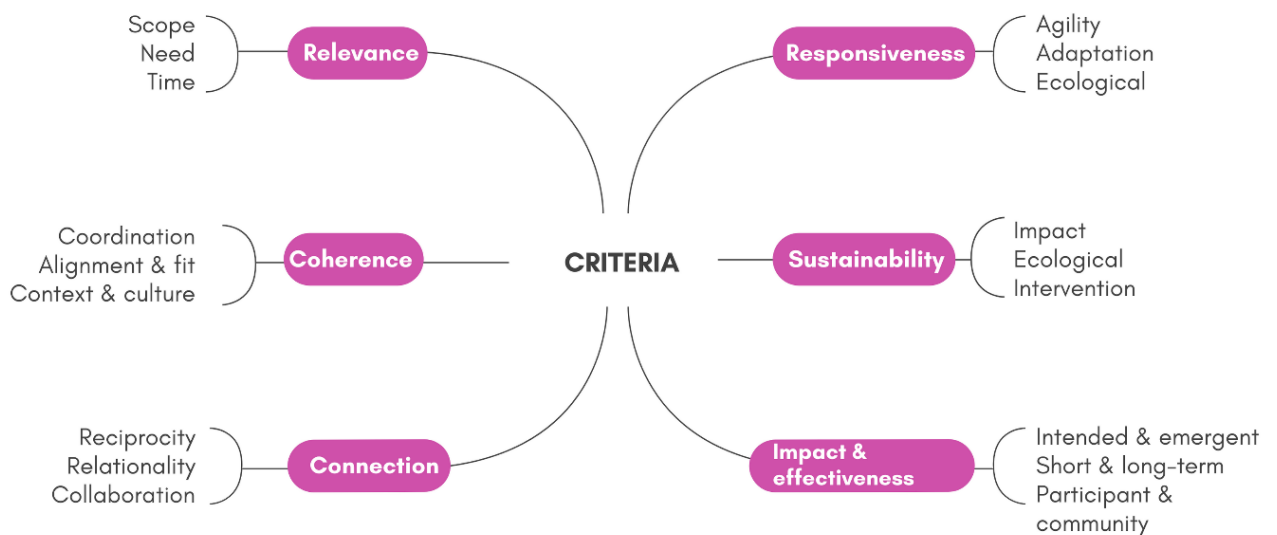
1. Relevance: The extent to which the programme and its scope (i.e., intervention design, intensity, reach) are relevant in light of identified needs within the country and target community, as well as the moment in time;
2. Coherence: How well the programme fits with and is aligned to the priorities of the country and the target community, as well as how co-ordinated or integrated the programme is with other, pre-existing programmes, services, and infrastructure. Coherence also includes the degree to which the programme fits with the context and culture of the community in which it operates;
3. Connection: The extent to which there is reciprocity between the programme and its community, as well as a collective sense of relationality, whereby the community is respected as knowers and active partners. Also included in this criterion is the extent or willingness to collaborate with other key stakeholders, especially the community in which the programme is situated;
4. Responsiveness: The degree to which the programme is able to exercise agility and appropriate adaptation to its context and environment, whether changes are substantial (e.g., a global pandemic) or relatively minor (e.g., gradual changes in community demographics). Furthermore, this criterion specifies the extent to which the programme is responsive to its ecological and climate context (e.g., whether it contributes to ecological degradation in its environment);
5. Sustainability: The maintainable longevity of the impacts and outcomes the programme produces, the ecological environment of the programme, and the programme itself;

6. Impact and effectiveness: Arguably the most important, this criterion assesses the extent to which the programme has produced desirable change in the short- and long term, including anticipated outcomes and impacts (i.e., specified in the ToC), as well as unintended or emergent changes. These changes should be observable in both the current programme participants and the community as a whole.

The criteria are shown in Figure 42.

Figure 42

Summary of SAMEL Evaluation Criteria



The Toolkit guides evaluators to assist programme teams in formulating definitions of criteria that are relevant to their programmes and contexts. The Toolkit encourages the use of systems concepts to create systemic definitions; e.g., what constitutes relevance can be considered from multiple perspectives.

Evaluation Questions. Evaluation questions are developed per criterion. As with the learning questions in the Monitoring & Learning chapter, the Toolkit proposes incorporating

systems concepts into the specification of evaluation questions (see Figure 43 for an example).

Figure 43

Toolkit Excerpt: Example of Evaluation Questions for Relevance Criterion

CRITERIA	EVALUATION QUESTIONS	SUB-QUESTIONS		
		<i>Systems & boundaries</i>	<i>Relationships</i>	<i>Perspectives</i>
RELEVANCE	Is the programme relevant to the needs of the community and the country?	How relevant are the programme boundaries? (intervention design?) How relevant are the target population boundaries in relation to context and need?	How relevant is the programme theory in relation to the context and need? How relevant are programme partnerships to the context and need?	How do participants and key stakeholders perceive the relevance of the programme?

Evaluation Methods. The Toolkit suggests a mixed-methods approach to evaluation, and provides a list of established methods that may be suitable. Three categories of methods are suggested: (1) qualitative approaches to capture open-ended and emergent phenomena (e.g., Outcome Harvesting, Most Significant Change, Photovoice), (2) qualitative approaches to capture anticipated phenomena or associative relationships (e.g., case study, positive deviance, contribution Analysis), and (3) quantitative approaches to measure anticipated phenomena and causal attribution (e.g., direct assessment, time series, dose-response analysis). To guide methodological choices, the Toolkit provides a resource that summarises examples of South African evaluations in ECD.

Evaluation Boundaries. When an evaluation is initiated for a programme, a new system is created. The evaluation system is distinct from other systems, such as the programme system or organisational system. It contains a different set of actors and expertise, and its efforts are directed towards a different goal. It needs new relationships to function, and should operate from a range of different perspectives (Gates, 2018). To set the scope and focus of the system, relevant and appropriate boundaries need to be established.

These boundaries can be drawn by, once again, turning to CSH (Ulrich, 1983). CSH can be used to determine and interrogate the boundaries of any system of interest, including the evaluation system.

Using CSH to critically reflect on the boundaries of the evaluation system may help to ensure that key normative assumptions, especially regarding stakeholder involvement, are adequately addressed. Towards this end, CSH can be used in the *ought* mode to guide decision-making on key boundary choices that will affect the evaluation's design.

The central component of CSH is boundary critique — the systematic questioning of boundary choices made for a particular system of interest, according to four categories: motivation, control, knowledge, and legitimacy (Ulrich & Reynolds, 2010). To unpack these categories, CSH provides three boundary questions per category: one question looks at social roles or stakeholders (the *who*); one looks at the specific concerns of those stakeholders (the *what*), and one addresses critical tensions between the two (Ulrich & Reynolds, 2010).

The answers to the 12 boundary questions describe how a system of interest (e.g., a programme, evaluation, policy, or problem) is framed and bounded, illustrating the values, norms, and assumptions that are at play. In the Toolkit excerpt shown in Figure 44, the CSH questions were adapted for framing an evaluation system.

The CSH questions, as applied to evaluation, help to prompt critical reflection on key issues relating to the evaluation's design:

1. The motivation questions frame the overall purpose and focus of the evaluation;
2. The control questions establish the key decision-makers, as well as how evaluation resources will be controlled;
3. The knowledge questions raise the issue of appropriate expertise and evaluation quality; and

- The legitimacy questions attend to the potential negative effects of the evaluation, together with methods to mitigate harm.

Figure 44

Toolkit Excerpt: Adapted CSH Questions for Evaluation Bounding

	SOCIAL ROLE / STAKEHOLDER	KEY ISSUE / STAKE	ADDRESSING THE TENSION BETWEEN SOCIAL ROLES & STAKES
MOTIVATION	<i>Beneficiaries:</i> Who ought to be the intended beneficiaries of the evaluation? Who ought to be the intended users of the evaluation?	<i>Purpose:</i> What ought to be the purpose of the evaluation?	<i>Measure of improvement:</i> What ought to be used to measure success?
CONTROL	<i>Decision maker:</i> Who ought to be in control of the evaluation?	<i>Resources:</i> What evaluation resources ought to be controlled by the decision maker?	<i>Decision environment:</i> What evaluation resources ought to be outside of the decision maker's control?
KNOWLEDGE	<i>Expert:</i> Who ought to be providing knowledge, skills and expertise to the evaluation?	<i>Expertise:</i> What knowledge, skills and expertise ought to be relevant for the evaluation?	<i>Guarantor:</i> What ought to be regarded as assurances of a successful evaluation design and implementation?
LEGITIMACY	<i>Witness:</i> Who ought to represent those who may be negatively affected by the evaluation (and who are excluded from the process)?	<i>Emancipation:</i> What opportunities ought to exist for those who are negatively affected by the evaluation to express their views?	<i>Worldview:</i> What space ought to be available for reconciling different worldviews regarding the evaluation (or its findings) among those directly involved and those excluded?

Note. Adapted from *Critical Systems Heuristics* by W. Ulrich & M. Reynolds, in M. Reynolds & S. Holwell (Eds.), 2010, *Systems Approaches to Managing Change: A Practical Guide* (p. 244). Copyright 2010 by The Open University.

The Toolkit advises that the CSH questions be answered twice at the beginning stages of an evaluation: once to brainstorm what *ought* to happen in the first stages of evaluation planning (the SAMEL workshop), and again to finalise what the evaluation plan actually *is*, once all planning decisions have been made (when the team is ready for an evaluation).

When the team is planning for an actual evaluation, what *is* and what *ought* to happen should be compared and critically appraised before the evaluation commences, to determine the difference and answer the following questions: What evaluation elements have been added or removed?; Are these decisions justifiable, and according to whom?

The two iterations of the evaluation framing will benefit the final evaluation plan by ensuring that the plan has been based on a thoughtful, value-setting exercise in which an ideal evaluation system has been laid out.

It is inevitable that the final evaluation system or plan will drift from the ideal one that was initially laid out in the SAMEL Evaluation workshop, due to various practical constraints. However, first engaging in mapping the ideal will ensure that the final system or plan is systematically bounded and, thus, designed to fit the ideal as closely as possible. This will help to protect the evaluation from morphing according to various changing influences or agendas. The Evaluation chapter provides guidance for answering the CSH questions.

Evaluation Readiness

The final session of the Evaluation workshop is aimed at establishing the programme's and team's readiness to undergo an evaluation. This work involves team discussions, intended to stimulate in-depth reflection on programme progress and organisational readiness. While checklists are often used to determine readiness for evaluation, open-ended questions may be more productive and constructive in surfacing and dealing with team members' assumptions and motivations that are likely to affect their participation in the evaluation.

Furthermore, discussions that promote the sharing of multiple perspectives and experiences may help to ensure a more effective and well-rounded evaluation plan. This could help foster a sense of ownership, thereby increasing team interest and participation in the evaluation, as well as practical use of the findings.

Programme Readiness. To determine whether the programme is ready to undergo evaluation, it is important to identify the programme's degree of maturation and the data collected. The Toolkit guides evaluators in conducting a reflective session with the team to discuss the programme's maturity, as well as considerations regarding time, personnel, and financial resources.

Team Readiness. As the SAMEL approach proposes a hybrid evaluation, the programme team and evaluation managers are likely to be significantly involved in the evaluation (albeit in different ways). As they should be amongst the primary users of the findings, it is important to address their initial perceptions and assumptions regarding the evaluation process (Patton, 2012). This may help the team think through the implications of evaluation and engender a better understanding of, and commitment to, the process.

In my consulting experience, I have often found that NPO teams feel unprepared for the logistical, operational, and strategic realities of evaluation. Even when equipped with a robust evaluation plan and a detailed planning process, if they have not experienced an evaluation before, it can be daunting and operationally disruptive.

It is thus advised that the Evaluation workshop open with team members anonymously sharing written keywords that represent what they fear most about evaluation, and what they find exciting. The Toolkit advises that the evaluator to read these out loud, noting similarities and acknowledging the team's sentiments (e.g., 'I can understand why that would be daunting'; 'Many people feel this way'; 'That's to be expected').

To attend to the team's feelings and expectations with regard to evaluation, I developed a boardgame for use in the Evaluation workshop. The evaluation boardgame is meant to introduce programme teams to the practical evaluation journey on a high level,

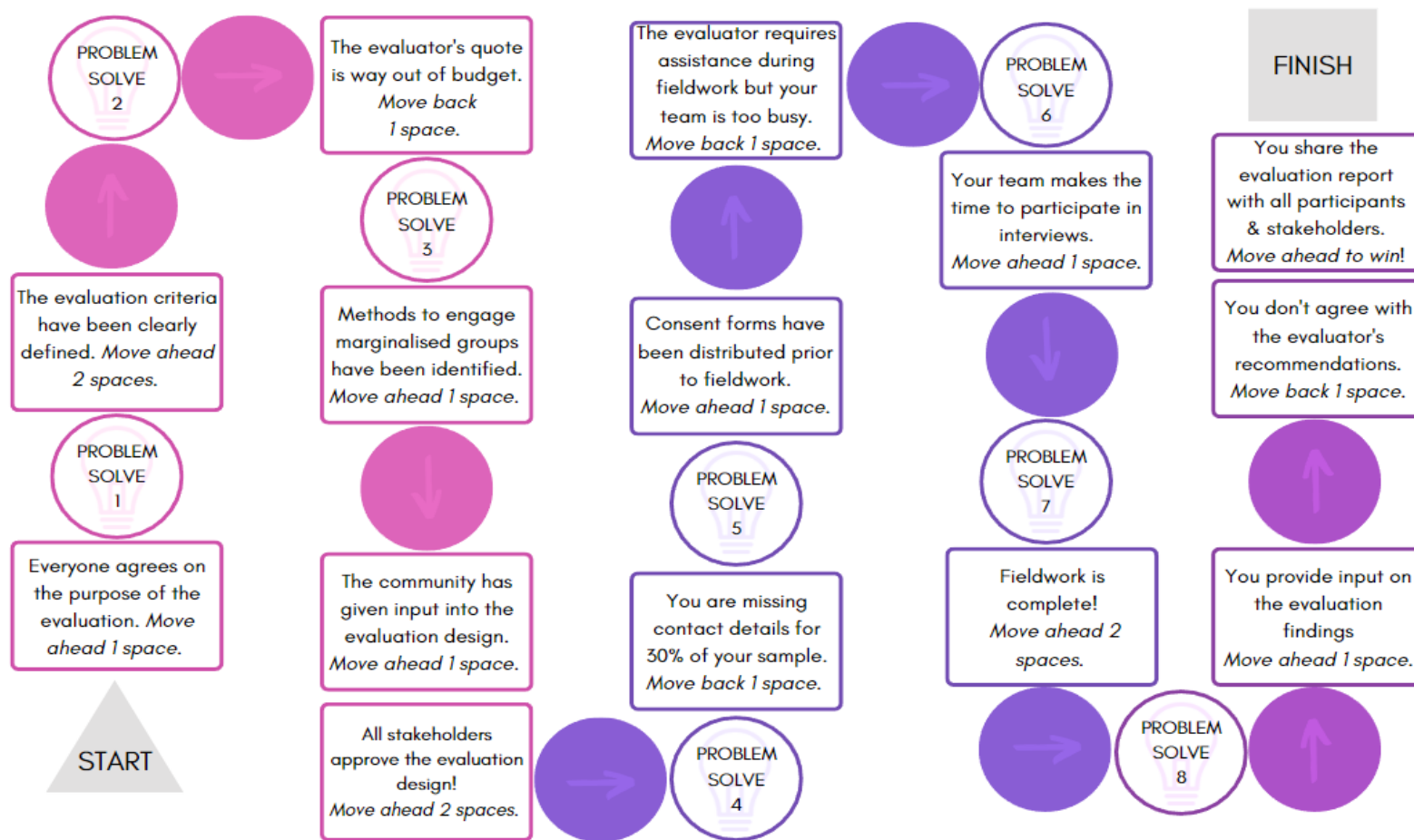
while the gamified nature of this exercise attempts to ease any potential anxiety or apprehension that may be associated with an evaluation.

The boardgame (see Figure 45) is played by splitting the team into smaller groups of two or more players. Each group member is given a gameboard and a token to move around the board (with moves determined by the throw of dice). The square blocks contain information about the fictional evaluation process that instructs the group member to move forward or backward on the board. This information alludes to generic experiences that the team members may face throughout the evaluation process, e.g., *“The evaluator’s quote is way out of budget. Move back 1 space.”*

The gameboard also contains ‘problem-solve’ spaces. If a group lands on one of these spaces, they select the corresponding problem-solve card, and the group engages in a problem-solving exercise. If the group can agree on a way forward, they can proceed on the gameboard. For example, *“You receive complaints from community members about the conduct of the fieldworkers. What do you do? Move ahead 2 spaces if you can agree on a solution.”* Most of the answers to the problem-solve cards would have already been discussed during the CSH boundary-setting session.

Figure 45

The SAMEL Evaluation Boardgame's Gameboard



This relatively short and simple game is intended to break the ice around evaluation and stimulate discussion of unanticipated issues that may crop up later. By engaging in fictional problem-solving, team members may feel more confident and ready to participate in an evaluation.

Following the boardgame, the Toolkit advises evaluators to conduct the readiness reflection session, focusing on the team and the programme. In addition to the discussion noted above on the programme's readiness, the session includes questions that centre on the team's perceptions of their competencies, expectations, values, feelings about the process, and dissemination plans, amongst others. A reflective worksheet is provided to help structure this session.

Expert Feedback on Evaluation Chapter

The original version of this chapter did not include the MAE principles or explicit evaluation design guidelines (including methodologies), but focused on CSH and the OECD DAC criteria (OECD, 2021). This was critiqued by Sarah and Alletta, who both felt that guidelines on design and methodology needed to be included, and that debate around the DAC criteria had not been addressed. Alletta noted the need for criteria to be operationalised and defined (which the original chapter did not attend to), and suggested the incorporation of meta-evaluation standards into the CSH exercise. Naledi and Alletta further suggested an evaluation reference group. Lastly, Sarah noted the need to deliberately include indigenous methodologies. I agreed with all these suggestions, and addressed them accordingly by:

1. Creating the section on evaluation design, including methodological suggestions and criteria definition;

2. Including the AfrEA evaluation guidelines (AfrEA, 2020) as part of the CSH section that deals with quality assurance (a worksheet was created to help teams reflect on the guidelines throughout the evaluation process);
3. Addressing the shortcomings of the DAC (OECD, 2021) criteria by incorporating Ofir's (2017a, 2017b, 2017c) and Patton's (2021) contributions, and offering the SAMEL criteria;
4. Suggesting that a steering committee or reference group be included in the CSH section that deals with the evaluation decision-makers; and
5. Incorporating the MAE principles into the Introduction chapter and the SAMEL criteria.

Furthermore, Sarah suggested that CSH would be *“more useful and insightful if done from multiple perspectives”*. Sarah held the view that the workshop was based on *“problematic trends in practice of having commissioners be decision-makers and evaluators working within that scope”*, and that additional perspectives should be incorporated by those who would be affected by the evaluation. While I agree that the rigour with which CSH is applied would be enhanced by including more programme stakeholders in this workshop, I disagree that it is problematic to have the programme team be decision-makers.

In my professional experience with small NPOs, it is more common for funders, rather than the NPOs, to act as evaluation commissioners. Often, terms of reference and evaluation designs are created by the funder, with little input into planning by the NPO teams. As will be discussed in the following chapter, this was indeed the case for the two test teams. For this reason, and for pragmatic considerations, I decided to focus the workshop on the core programme team.

Sarah also felt that the game was too basic, and that it trivialised evaluation issues. Sarah thought case studies would be more useful. I opted to include the game despite this feedback, as I felt that evaluation cases would be too academic. I am also of the view that in-person workshops benefit from higher-energy activities.

The next section provides a summary of the experts' feedback on the entire SAMEL Toolkit.

Expert Review of the SAMEL Toolkit: Overall Comments

The experts' consolidated high-level feedback regarding the overall quality and feasibility of the Toolkit is provided below.

All reviewers indicated that the SAMEL Toolkit does meet a need. They noted that the Toolkit fills an important gap in terms of making systems thinking accessible and practically applicable for evaluators, and that it would assist in mainstreaming systems thinking in evaluation. Alletta noted the gap that the Toolkit was designed to fill, saying:

"There is a paucity of guidelines/tools to assist a non-academic evaluator to apply systems/complexity concepts to M&E work. This toolkit makes an important contribution. All steps and many — if not all — of the tools/worksheets, as well as the workshop outlines, are useful in this regard."

The reviewers noted the Toolkit's ability to expand evaluator competencies, and to support systems thinking capabilities amongst both evaluators and NPO teams. Alletta said that, if properly facilitated, this could also enable greater evaluative development within NPOs. Alletta and Sarah both felt that the Toolkit is focused on evaluator facilitation, and not internal facilitation by NPO team members, with Alletta saying that an NPO team would struggle to facilitate it.

With regard to evaluators' MEL competencies, all reviewers felt that the Toolkit contributes valuable additions to an evaluator's skillset, particularly the Scoping and Focusing chapters. These chapters, and the systems thinking tools contained therein, appeared to be particularly well received by the reviewers. Alletta noted that the Monitoring & Learning chapter would also be particularly useful in enabling learning in organisations, stating:

"Chapter 4 — Develop a monitoring plan and learning strategy — is very, very valuable, especially the guidance on developing a learning strategy (learning is under-emphasised in a number of how-to develop an M&E framework/system guides)."

All reviewers mentioned a need for more practical examples throughout the Toolkit, with Alletta noting that much of the systems thinking content seemed "dense" or "theoretical". Alletta felt that an evaluator might find it difficult to conduct the SAMEL workshops without clearer guidance or support. However, Sarah, the expert most experienced in systems thinking, described the Toolkit's language as "clear, engaging and applicable". These divergent views demonstrate the need for the language around systems thinking to be made more user-friendly for newcomers, as well as the need for clearer guidelines for evaluators (both of which were addressed as detailed in the preceding sections).

In terms of feasibility, Alletta and Naledi both felt that the full SAMEL process might not be feasible for all NPOs, particularly in terms of time and resources. Naledi noted that workshop time would have to be considered very carefully, and Alletta wondered whether a lighter version of the Toolkit should be made available for less-resourced NPOs. Sarah noted that the Toolkit was "very feasible", but noted having a limited perspective of the South African NPO setting.

Sarah emphasised two important theoretical issues. First, Sarah noted that the Toolkit rests on the assumption that a programme and MEL system should be designed by the NPO team, saying:

“From an ethical and political perspective, the toolkit seems to assume that the NPO team and evaluators are the right people to design the programme and the MEL system in consultation with others. This greatly depends on who is and who is not on the team.”

While I agree with Sarah’s argument, I do not believe that there is an alternative that is pragmatically feasible. To the best of my knowledge, it is very rare for funders to support programme- and MEL design driven by communities or other groups exclusive of NPOs and evaluators. Instead, the Toolkit proposes a stakeholder analysis in which NPO teams can decide who works alongside them in the SAMEL workshops. This does, however, place the power solely in the hands of the NPO, which is another pragmatic trade-off that requires testing and refinement.

Sarah’s second major critique was that the Toolkit takes a conventional approach to MEL:

“The toolkit is based on conventional logic of social problem-solving based on diagnosing a problem, designing and implementing an intervention, and MEL. It’s infusing systems thinking into that, rather than challenging it.”

This was an accurate critique, and one I weighed up continuously throughout the development phase. As noted in the previous chapter, I ultimately decided to take a more conventional approach, so that the Toolkit would fit in with the current way most NPOs and funders work. I felt that it would not be practical or useful to design a completely innovative

process that many NPOs would not have the freedom to enact, due to their funders' requirements.

Chapter Conclusion

This chapter described the content of the SAMEL Toolkit, which is structured into five chapters. I explained the systems thinking tools used per chapter (and how these integrate with the systems concepts *systems, boundaries, relationships, and perspectives*), and how they should be applied by evaluators.

In summary, the Introduction chapter provides an overview of the systems concepts, how these relate to MEL, and the rationale behind the Toolkit. The Scoping chapter introduces tools for conducting a situation analysis (such as the SEM, CLDs, and system archetypes) and developing the foundation of the programme strategy. The Focusing chapter builds on this foundation by developing a programme theory, consisting of a ToC and a ToA. The Monitoring & Learning chapter guides the development of a theory-based monitoring system, and facilitates continuous learning through reflective practices. Finally, the Evaluation chapter is aimed at enhancing evaluation readiness through the development of an ideal evaluation plan.

Expert feedback was instrumental in refining the Toolkit, highlighting both strengths and areas for improvement. The majority of suggestions by the reviewers were incorporated into the final design, as detailed in this chapter of the thesis.

Reviewers acknowledged the Toolkit's role in addressing a significant gap in MEL by making systems thinking accessible to evaluators. However, they emphasised the need for more practical examples to aid comprehension and application of the theoretical content. Concerns were raised about the feasibility of the SAMEL process for NPOs with limited

resources, and some critiques pointed out the Toolkit's reliance on conventional problem-solving logic and the NPO team's perspectives, which might limit its transformative potential.

The next chapter describes the SAMEL test, detailing the Toolkit's implementation in real-world settings.

Chapter Eight: Testing of the SAMEL Toolkit

This chapter presents a description of testing the SAMEL Toolkit with the EYEN and EduSpark teams. This chapter draws on the following data sources to develop a case description (Yin, 2018):

1. A total of 10 transcripts of focus group discussions in each workshop conducted for both test cases to elicit the NPO teams' opinions of the workshops;
2. Observation notes from each workshop in both tests, which recorded the day's activities, as well as how both the evaluator and teams engaged within each workshop session (a total of 10 observations);
3. My reflexive journal for each workshop that I facilitated, in which I recorded my experiences of and reactions on the day (five sets of journal entries);
4. Informal check-ins with Heather before, during, and after the workshops that she facilitated, to gauge her experience of facilitating the Toolkit content (a total of seven recorded conversations);
5. Final debriefing interviews with each organisation to understand the aftermath of the workshops (a total of two interviews); and
6. A final debriefing interview with Heather to reflect on her experience of the entire test process and the Toolkit.

In order to describe how the SAMEL Toolkit was facilitated and engaged with across the workshops, the data were consolidated using a descriptive framework (Yin, 2018). The SAMEL workshops were used as descriptive topics (i.e., Scoping, Focusing — ToC, Focusing — ToA, Monitoring & Learning, and Evaluation), and each NPO case was used as a sub-topic. After initial immersion in the data, the data were sorted into these topics and then consolidated to develop a chronological narrative. This chapter is thus

descriptive only; it provides the reader with an evidence-based account of what occurred in each test, to create an understanding the practical challenges of, and initial reactions to, the Toolkit. An interpretive lens was applied to these findings through reflexive thematic analysis. The reflexive thematic analysis is presented in Chapter Nine.

A key aim of this phase of the research was to test the SAMEL Toolkit in real-life settings and answer the following questions: How feasible is a systems thinking approach to MEL development? How did the workshop participants and the evaluators (workshop facilitators) experience the SAMEL Toolkit as a systems thinking approach to MEL? What are the challenges in and advantages of integrating systems thinking into MEL?

This chapter unfolds chronologically, beginning with the planning of the test phase, followed by a discussion of each workshop conducted for each test case. Each specific workshop is described for both test cases, to highlight similarities and differences in the facilitation of the workshops and the NPO teams' experiences and impressions. This is done to give the reader insight into how the workshops were delivered, how the teams responded to the SAMEL content, and what constraints influenced the process as a whole.

Test Planning and Team Commitment

“Particularly now, as it's crunch time for us in terms of where we are currently, do we have the time to invest in this? Can we afford time out of the office, a whole day in this case? Will it be of any benefit to us?” (EduSpark CEO, Scoping focus group)

In the extract above, EduSpark's CEO questions the practicality of attending SAMEL workshops due to the considerable time required. While they went on to answer with *“definitely”*, the challenge of setting aside time for the SAMEL test was a critical consideration in both the planning stages and the execution of the workshops.

I initially presented the SAMEL methodology to EduSpark and EYEN representatives. At EduSpark, the CEO and two managers were my main contacts, while at EYEN, I worked with the MEL officer. Both organisations were eager to join the test, as they were aiming to revise their programme designs and MEL systems amid budget limitations. EYEN's MEL officer, in particular, sought guidance due to a lack of MEL experience. In a focus group, the MEL officer expressed gratitude for the support and knowledge shared through the SAMEL test, emphasising its value in enhancing their work:

“I think everybody around the table is very grateful for your time and yourself and Heather, and the efforts that you're putting into this process with us, and having to deal with all our personalities [laughs]. I do feel like the process is building on to something bigger, and we will use the skills and the information that we've gained through this process to apply it across the board...” (EYEN MEL manager, ToA focus group)

During the planning stage, scheduling the workshops and ensuring team attendance were challenging in both settings, due to the need for a five-day commitment. One full day was needed per workshop: Scoping, Focusing — ToC, Focusing — ToA, Monitoring & Learning, and Evaluation.

I encouraged inclusive participation from each organisation, aiming to include a range of staff involved in management and implementation. However, operational demands made this difficult. Many team members, particularly those involved in fieldwork, were unavailable to participate in the workshops, as I was told that they could not be pulled out of the field or the office. Time was a significant constraint, as highlighted by EYEN's MEL manager:

“Our team is very scattered most of the time because of the nature and work that we do. Not everybody's in the office all the time. So, getting everybody in one place was a mammoth mission. But we managed. Thank you, Gail, for making it compulsory,

otherwise people wouldn't have shown up.” (EYEN MEL manager, Evaluation focus group)

EduSpark's CEO also acknowledged the intense time commitment required for the workshops in a debriefing interview:

“I mean, the process was very, it was very intense. It took us out, you know, a full day out of the office, or several days, which we cannot really afford, but I think it was worth the while.” (EduSpark CEO, debriefing interview)

My observations during the workshops confirmed that team members often had to multitask due to time constraints, balancing workshop participation with ongoing work responsibilities. This resulted in frequent interruptions and absences as they attended to emails, phone calls, or urgent issues. Additionally, tasks assigned for completion outside of the workshops were seldom done, and EYEN workshops had to be rescheduled several times.

Heather, in her final debriefing interview, pointed out the difficulty of finding a suitable time for such strategic initiatives. Fluctuating operational demands, such as reporting periods and external commitments, significantly limited availability:

“It also depends on the time of the year, right? If it's a busy reporting period time. Because like, for them, for June, there were three weeks in June that they just couldn't because they had funder visits, or they had reports due, or whatever it was. Now I'm thinking, like, November, December, January, February is also like a write off because of closing off, reporting, and starting the year. So, I don't know when a good time of the year would be.” (Heather, debriefing interview)

Scoping Workshop

“We don't just want to save that one, we want to save the world!” (EYEN CEO, Scoping focus group)

In this extract, the CEO of EYEN expressed the NPOs' ambitious and compassionate drive to enact widespread impact. This ambition underscores the Scoping workshop's aim of enhancing teams' understanding of their specific contexts, thereby enabling a more focused and impactful programme design. However, the broad focus of NPOs often complicates the process of narrowing their objectives.

To remind the reader, the Scoping workshop is designed to immerse teams in their specific context through a situation analysis. This situation analysis lays the groundwork for defining a focused problem statement and articulating the organisation's mission, vision, and values. To aid this analysis, the SAMEL Toolkit introduces a variety of systems thinking tools: the SEM (Bronfenbrenner, 1979), CLDs and system archetypes (Senge, 1990), actor mapping (Gopalakrishnan & Clarke, 2015), and SSM/CATWOE (Checkland, 1981) (presented in Chapter Seven).

In both tests, Heather and I completed one aspect of the situation analysis (background information on ECD) prior to the workshop commencing. I also drew up the SEM (Bronfenbrenner, 1979) using relevant examples. This was necessary due to time constraints and because we both individually felt that the situation analysis would be enhanced by collaborating with the teams to get their on-the-ground experiences and insights. This was particularly important in the EduSpark test, which I facilitated, as the management team had decided to invite a number of ELP principals (past and present) to the workshop.

The Scoping chapter has the most systems thinking content of the Toolkit. Heather, being completely new to systems thinking, found the chapter to be academically demanding. Heather made a light-hearted comment about this during a pre-workshop check-in, indicating a lack of confidence in her abilities:

“It's making me feel very academically inadequate!” [laughs]. (Heather, pre-Scoping workshop check-in)

Subsequent sections detail how each Scoping tool was facilitated and how the evaluators and teams interacted with them, underscoring the practical application and engagement with the systems thinking approach outlined in the Toolkit.

Tool 1: Evidence

The SAMEL Toolkit guidelines suggest starting the Scoping workshop with an overview of the situation being addressed by the organisations, using published evidence.

Heather and I presented an overview of the current state of early learning, focusing on children's access to ELPs. This was a straightforward process for both of us, although the depth of our presentations differed. Heather used three Toolkit-provided data sources for a broad historical perspective on ELP access in South Africa. My approach, benefiting from deeper experience in the ECD sector, involved more recent data, particularly relevant to EduSpark's operational region.

Both NPO teams showed keen interest in the data, finding it valuable for understanding their reach and impact. The EduSpark team appreciated insights into regional ELP support, as it helped them to see how they were contributing to their communities. The team also felt that the information affirmed their chosen intervention strategy, highlighting that there is a significant need for the services they offer.

Tools 2 and 3: The SEM and Actor Map

Following a review of existing evidence, the Toolkit suggests addressing the systems and boundaries of the situation analysis using the SEM (Bronfenbrenner, 1979) and actor maps (Gopalakrishnan & Clarke, 2015). Adaptations were necessary for both tools to fit the workshop's dynamics and participant availability.

I had to adapt the EduSpark workshop agenda due to team delays causing a late start and the inclusion of ELP principals, who could only join the workshop in the morning. I had to skip the SEM discussion and reschedule the actor mapping for later in the day. This shift prioritised the principals' perspectives for the CATWOE session (discussed later), as well as a discussion of the situation's problems and causes. Engagement fluctuated, particularly before and after lunch, but, ultimately, the session was productive, focusing on identifying key actors and prioritising urgent relationships.

Heather facilitated the SEM and actor mapping exercises, and showed a good understanding of the underlying concepts. I observed that the sessions were facilitated clearly, and were met with high interest and engagement from the team. However, I noted that Heather's written materials were difficult to see from the back of the room, as her slides were dense with writing and her handwriting on the whiteboard was small (see Figure 46).

The sessions provided valuable insights, with team members expressing appreciation for the comprehensive overview offered by the SEM, with one commenting:

"It's very powerful to put it together and see the interconnectivity between the levels."

(EYEN team member, observation)

In the EYEN focus group, there was widespread appreciation for the SEM and its ability to help the team understand the ECD sector in a comprehensive yet digestible way. Heather echoed the team's sentiments in her final debriefing interview, saying that the SEM session was a standout for her.

In EYEN's actor mapping session, challenges arose in identifying the core of the actor map. I observed Heather and the team grapple with the notion of influence as they began the session with children at the core of the map, which they then changed to ELPs, their direct

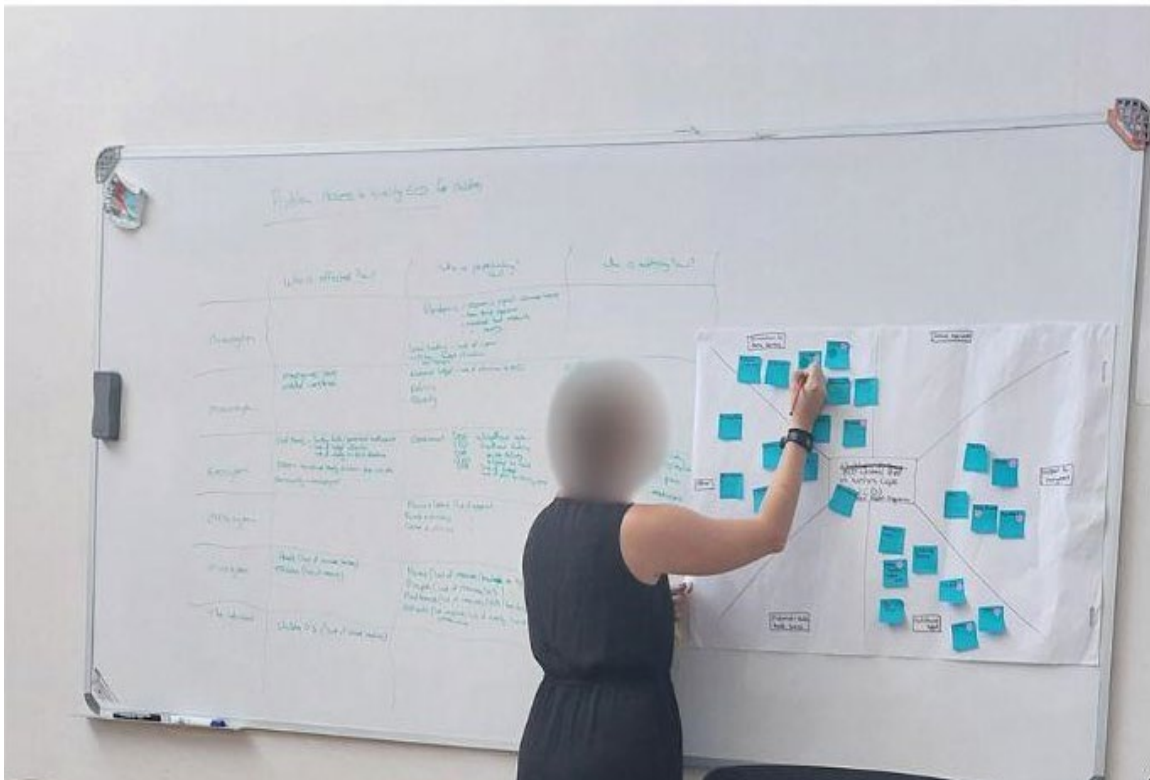
contact participants. I observed that the team was able to easily name actors, and that the team's CEO was particularly skilled in identifying opportunities using the map.

Reflecting on the actor maps in her final interview, Heather noted the teams' enthusiasm, but described her own hesitation as follows:

"From an M&E framework perspective, it's not that useful. But I think, from their perspective, it was really useful. I think they found it really good to think this through and identify these things. And I think facilitating that process for them was important, not necessarily as important as the output." (Heather, final interview)

Figure 46

Heather Facilitating EYEN's Actor Map



Tool 4: Causal Loop Diagrams and System Archetypes

The SAMEL Toolkit recommends using CLDs and system archetypes (Senge, 1990) to explore non-linear relationships that contribute to the situation being investigated.

Heather struggled to grasp these concepts. In a check-in call before the Scoping workshop, she communicated that she was not comfortable with the concepts, saying, “*They did not click*”. This discomfort did not ease during the workshop. In my final interview with Heather, I asked what she struggled with the most. In a light-hearted but seemingly fed-up response, she said:

“Those flippin’ feedback loops [laughs], feedback, archetypes, and all of that stuff, because they were completely new concepts. It’s, I don’t know, it did not click with me. And I didn’t necessarily have the time to research it enough so that I could make it click with me.” (Heather, final interview)

My own experience mirrored Heather’s to some extent. I found the underlying concepts challenging to internalise while developing the Toolkit, and I needed time to feel comfortable with them. Despite a better understanding of the concepts, I was still anxious about their practical application in the workshop, and lamented in my journal: “*CLDs — my biggest worry!*”

In the EYEN workshop, Heather introduced CLDs using a video, and engaged the team in identifying causes for the identified problem (the lack of child access to quality ELPs). Heather’s own uncertainty led to a brief and somewhat superficial discussion of identifying the causes of the problem using these tools. In a check-in following the Scoping workshop, Heather reflected:

“If I had been more confident, I would have done more with it. If that makes sense. I would have done more examples and whatever on the board.” (Heather, post-workshop check-in)

Despite a somewhat lacklustre introduction to CLDs, I observed interest from the EYEN team. They engaged well with the concept of non-linear relationships, and identified two reinforcing relationships, showing that the introduction of CLDs had created some level of understanding. The discussion of system archetypes was limited to a presentation of examples from the Toolkit chapter. I observed Heather advising the team that *“we are just planting a seed”*, seemingly unwilling to delve deeper, due to her discomfort with the concepts.

At the end of the workshop process, one EYEN team member, who held a doctorate in science, reflected on the concept of feedback as a key takeaway from the workshop experience:

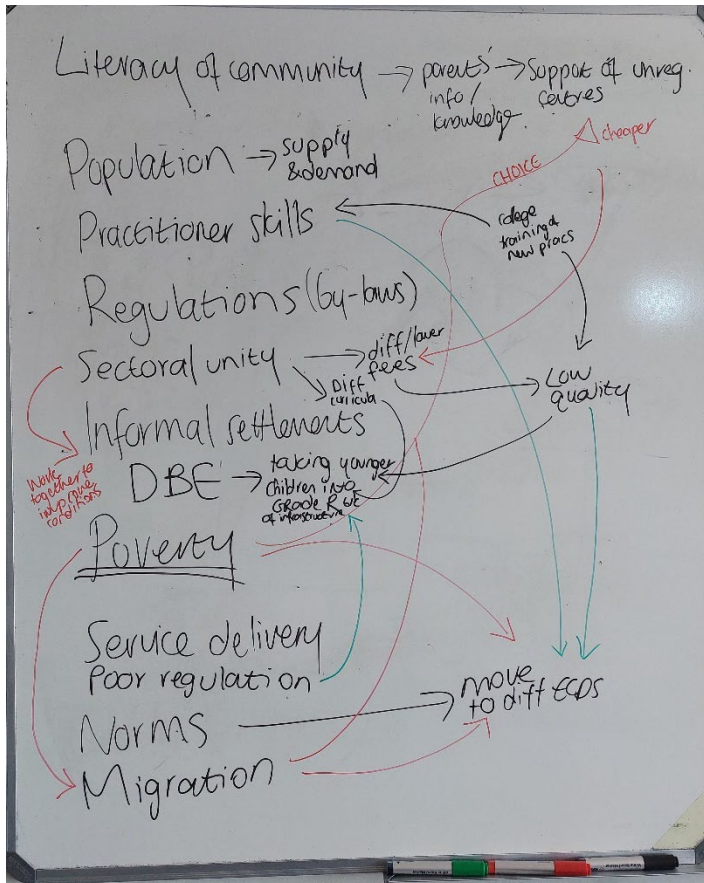
“So, that's one thing that stood out, and I keep thinking about how one element feeds into the other element, and how they affect each other.” (EYEN team member, Evaluation workshop focus group)

In the EduSpark workshop, I anticipated difficulties in explaining CLDs to an uninitiated audience, particularly as the group was large due to the inclusion of principals, and I was stressed during the introduction of the concept to the group. Unsurprisingly, I observed significant hesitation after I had suggested that we use CLDs to understand what might be going wrong in the community's ECD context (the situation). Cognisant of time limitations to the principals' attendance, I did not want to spend additional time getting everyone to understand CLDs. Instead, I abandoned CLDs and simply asked the group to talk freely about what was going wrong in the local ECD community. Using a whiteboard, I first linearly listed

causes, and then linked them together as the discussions expanded, highlighting relationships between them (see Figure 47).

Figure 47

Amended CLD Session: Listing and Linking Causes



Despite deviating significantly from the Toolkit, this approach still resulted in a productive session, with the EduSpark team gaining insights into the complexities of their operational context. The EduSpark focus group appreciated the opportunity to discuss real-world challenges from the perspective of the principals, obtaining rich insights into the community's needs instead of relying on their own assumptions. EduSpark's CEO highlighted this in the following focus group extract:

“We talk about it... 'just by the way', but here was an acknowledgement of it, an example by one of the principals, and, currently, what is actually happening. That's a highlight to me, in terms of, you know, what the realities are on the ground for these principals.” (EduSpark CEO, focus group)

Tool 5: CATWOE and Root Definitions

Borrowed from SSM (Checkland, 1981), CATWOE and root definitions are intended to help teams understand the various ways in which different actors may perceive the situation of interest.

ELP principals attended the EduSpark session, and the CATWOE activity was completed in groups. To aid comprehension, I used the alternative acronym *BATWOE* (replacing ‘customers’ with ‘beneficiaries’).

The observation and my journal reflections indicated good initial engagement and understanding, as I reflected that the prompts *“seemed to nudge good discussion among the team in terms of their own perspectives”*. However, I went on to describe the discussion following the group work as lacking depth and diversity in perspectives. While the actual exercise appeared to go well, the discussion afterwards was not particularly insightful. I speculated that the uniformity of stakeholder representation and possibly the timing of the session (too early in the day) were possible reasons for the subdued participation.

Figure 48

Me Facilitating the CATWOE Discussion



Heather took a different approach in the EYEN workshop, due to the absence of various stakeholders. In an empathic exercise, she encouraged the team to consider the situation from the perspective of different actors. I observed some struggles due to an overlap in actors and owners, while identifying constraints came easily to the team. In my observation notes, I wrote that Heather “*struggles to get the root definitions right at first*”.

According to the focus group data, the CATWOE exercise, combined with the SEM and actor mapping, provided powerful strategic insights for the EYEN team. While the SEM helped them appreciate where EYEN fits into the broader ECD picture, the actor mapping and CATWOE sessions helped them understand what is within their reach and what role they could play effectively in the sector in relation to other stakeholders. An ongoing joke within the team humorously summarised this clarity, that EYEN is not the “*main character*” in the ECD system, but rather an “*extra*” supporting a larger narrative. In the final workshop, two team members reflected on this joke as a key takeaway from the Scoping workshop:

Gail: “I think the first workshop was on the ecosystem.” [laughter]

Zara: *"I was waiting for this! This was the best part for Gail!"*

Gail: *"Traumatic but... I learned a lot because it's understanding systemically where we are relevant and where we make a difference. Your focus or your influence versus your area of concern. And to actually, you know, stick to your stuff. And I do have a problem with focus because I want to save the whole world. And, you know, I had to accept..."*

Zara: *"That you are only the extra!"* [laughter]

This anecdote, while light-hearted, underscored a profound learning moment for the team: understanding their specific place and function within the broader ecosystem. Heather, however, remained sceptical about the practical benefits of CATWOE. In my final interview with her, she said the following about the root definitions:

"I don't know how much benefit they add to, like, a frameworking process" (Heather, final debriefing interview).

Programme Scope

The Scoping workshop aimed to establish a programme scope by developing a clear problem statement and articulating the programme's mission, vision, and team values. This is used to guide resultant decision-making in the development of the programme theory. This constituted the final workshop session for the day, with both test cases' sessions taking place in the afternoon.

In both the EduSpark and EYEN workshops, while the teams had existing statements and were familiar with the concepts, they were observed to encounter difficulties in focusing and clearly defining their problem statement, vision, and mission. For example, I noted in my journal about the EduSpark team: *"There seemed to be some confusion amongst the team about what they are trying to achieve."*

I observed a similar hesitancy amongst the EYEN team, observing that the CEO was referring to a printout of an old logframe during this discussion. There appeared to be a sense of pressure to develop highly impactful statements, as can be seen in the focus group exchange below:

Gail (CEO): *"For me, the challenge came with vision and mission. It's not the easiest thing to articulate. I mean, you know what it is on an instinctual level. And you know what it's not. There's still a bit of work that is needed at that level."*

Zara: *"It's just our first draft, Gail."*

Gail: *"... We're trying to condense something that's not, and I think that's the challenge in articulating the mission. Because it's big."*

Zara: *"It's bigger than just..."*

Gail: *"We do a lot. And so, how do you just condense it in two sentences, where it's not able to be contained within that?"*

Despite these challenges, the process was seen as beneficial, especially for EduSpark. In the focus group, they described how engagement with ELP principals during the problems-and-causes session aided in refining their strategic statements. Sentiments expressed indicated that this process marked a departure from their usual practice of recycling outdated strategic statements from policy documents and *"parking"* strategic ideas without deliberately documenting them. One EduSpark team member appreciated being stimulated to *"naturally"* develop these statements in a strategic discussion forum, compared to isolated, referential drafting based on old organisational documents.

Focusing Workshops

“We're sitting with information... Bits there, bits there, some in more detail than others, and a lot in our heads. And we haven't gone through a structured process of unpacking all of that.” (EduSpark CEO, ToC focus group)

The extract above illustrates why the SAMEL Toolkit incorporates programme theory into the MEL framework: to effect collective deliberation on the programme's design, including its impact-related objectives and implementation process.

SAMEL Focusing takes place in two workshops that deal with programme theory, one each for ToC and ToA. For both test cases, each workshop was afforded a full day. The sections that follow describe each of these workshops.

ToC Workshop

The SAMEL ToC encompasses programme impact, outcomes, realist mechanisms, contextual conditions, and feedback loops. The full-day workshop introduces systems thinking through non-linear relationships and feedback loops, alongside realist evaluation mechanisms and contextual conditions (Pawson & Tilley, 1997) (using Kurt Lewin's [1936] Life Spaces for the identification of contextual conditions).

Interventions. While not a systems thinking tool, the workshop sessions on intervention definition must be discussed, as they revealed key learnings for theory development.

When developing the Toolkit, I expected the identification of programme interventions for the ToC to be a straightforward process. However, the teams in both test cases struggled to concisely describe the particular interventions that make up their programmes. Time-consuming debate was observed amongst both teams around what was

currently implemented, what it should be called, and if something different was needed. For example, for the EYEN ToC workshop, I made the following note of an observation:

“It’s not easy to simply list interventions. [Programme manager]’s answer is long and confusing. A lot of ‘why’... Lots of overlap and inter-connection between the interventions... A lot of problem-definition thoughts coming through here.”

These discussions appeared to be significantly influenced by contextual challenges within the programmes’ operating environments. These challenges surfaced as the teams reflected on and debated their interventions, with a particular focus on policy. South Africa's ECD policy prioritises ELP registration with the Department of Basic Education to gain access to government subsidies, creating pressure to focus on registration-enabling interventions (such as enhancing infrastructure, providing first-aid training, and training in administrative skills). However, the rigid registration requirements, like land ownership, pose dilemmas for organisations offering ELPs that are unable to meet them.

This became a sticking point in both test cases as teams grappled with the issue of government registration and whether there were more important outcomes to focus on, as pointed out by EYEN’s CEO:

“Do you know what? We can register better than we did ten years ago, but so what? [laughs] Our registration is fantastic — we can press a button and we register. But what’s it doing?... Because we’re measuring a lot of the processes; we’re measuring our registration, we’re measuring infrastructure, we’re measuring ‘Do they have the qualifications’, but we’re not going deep in terms of what’s really going to change or shift what we need to shift.” (EYEN CEO, ToC focus group)

In addition to policy pressure, discussions highlighted the role played by donors, as NPOs often design their programmes to align with grant proposals. In the EYEN workshop, I

observed Heather encouraging the team to be internally driven and not funder-driven in determining their strategy, as captured in the exchange below between Heather and the EYEN CEO:

Heather: *“Being internally driven and knowing what you are accountable for means that you attract the right funders, after the fact...”*

Gail: *“Yes, ... and that's exactly why this is critical. Because if you're not clear on this, you're going to grab everything that comes your way. And that's the challenge.”* (EYEN ToC observation notes)

As such, the types of interventions that form part of the ToC may end up being incoherent, as the NPO, in Gail's words, grabbed all funding that came its way and is now committed to delivering a range of donor-driven interventions. This was echoed in the EduSpark workshop:

“We've been implementing a very haphazard ECD model that was very much wide stretch and actually donor-led.” (EduSpark team member, focus group)

Despite these challenges, the process was valued for fostering collective engagement in defining the respective programme models, as doing so created a sense of ownership over the programme and its objectives, a sentiment that was aired in both ToC focus group sessions. For example:

“It's always great to get everyone's perspectives on the programme and scope together. I think, you know, in order for them to take ownership, they need to have ... a piece of skin in the game. So, I think everybody feeding into the model development is worthwhile.” (EYEN managing director, focus group)

“The fact that we could be in agreement, that we spoke about and debated against one another, around the issues that were mentioned and so on, leads to ultimate, not just agreement, but ownership.” (EduSpark CEO, focus group)

Mechanisms and Outcomes. Following the definition of the programme interventions and desired long-term impact, the Toolkit suggests that teams be encouraged to define observable programme outcomes. To enable teams to distinguish between realist mechanisms and outcomes, it is suggested that outcomes be viewed as externalised and observable change, whereas mechanisms relate to internalised change. This was a distinction I gleaned from the work of Dalkin et al. (2015).

Both teams had no difficulty identifying outcomes, evidenced by discussions around visible changes. This process was observed to be straightforward, with both teams readily coming up with ideas.

As with interventions, the role of stakeholders surfaced during outcome discussions. The EYEN team noted pressure from donors to measure child outcomes, over which the two NPOs have little control, as neither works directly with children. For example, I noted that the EYEN managing director questioned where the assessment of child outcomes came into the ToC, with Heather noting that the team cannot be accountable for child-level change. I observed that Heather was able to maintain the session scope with skill, respectfully but assertively probing the team for clarity and refinement.

Facilitating discussions around mechanisms proved more challenging. Both Heather and I understood the concept theoretically, but found its practical application difficult. I described my confusion in my journal, noting that I focused on internal cognitive changes (reasoning) in identifying mechanisms, which limited the possibility of identifying non-person-centred change (such as ELP infrastructural improvements). I also was unsure about the best

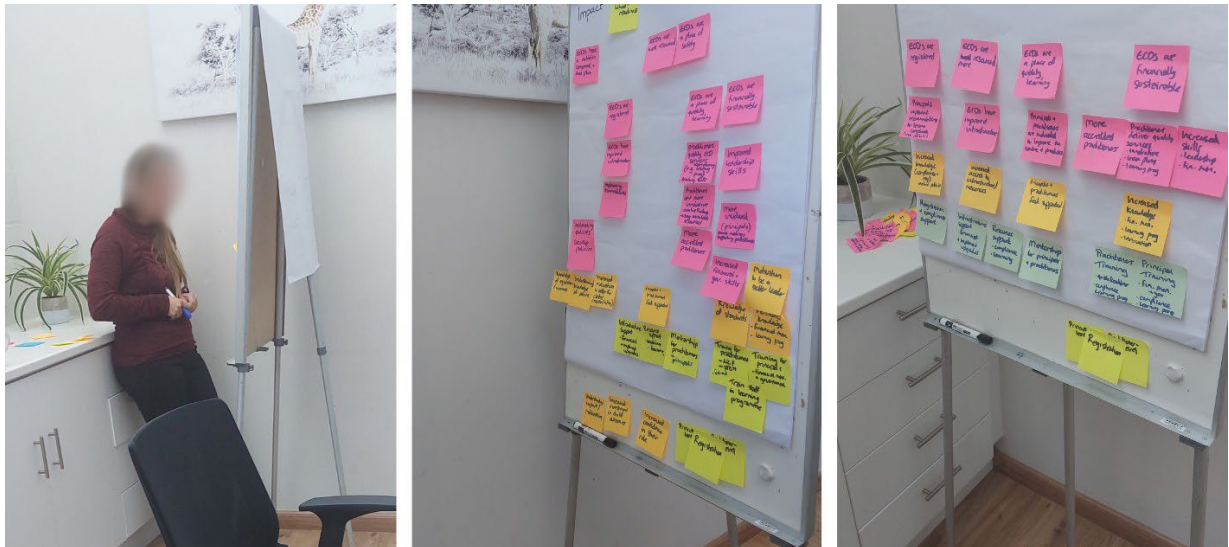
way to incorporate resources into the ToC. My confusion appeared to affect the session, as the observer recorded it as slow in pace and *“feeling heavy”*, with limited comprehension and engagement amongst the EduSpark team. A similar drop in energy was observed in the EYEN workshop. While I observed Heather explaining the concept clearly to the team, she remarked: *“I see a lot of confused faces.”*

As we grappled with the practical application of new methodologies, both Heather and I showed a tendency to revert to the more familiar linear methodologies. In my journal, I noted that, as time dwindled in the EduSpark workshop, it felt more efficient to keep the conversation linear. In the EYEN ToC workshop, I observed that Heather seemed *“overwhelmed by the messiness”* of mechanisms and outcomes, and spent the lunch break neatening the diagram being worked on to create a more linear-looking framework (see Figure 49). Similarly, the EYEN CEO expressed feeling overwhelmed by the *“chaos”* of a ToC, communicating a need to cut through complexity and hone in on a focus area, and showing appreciation for Heather’s linear and neat diagram:

“Earlier on, when you saw all that different colours all over, it was just one big, like a lot of chaos. And when you look at that, you're overwhelmed. And now, when you get to this [rearranged diagram], it's actually more structured, and you can process better, and you can actually approach it in a different way. Because, like, when you're looking at too much, it's just like, 'Oh, everything's important!'. And then you focus on everything, and you want to be everything to everyone, but you realise, 'Let's reduce our area of significance, of inputs, because we can make that much of a bigger difference’” (EYEN CEO, ToC focus group)

Figure 49

Heather Neatens the EYEN ToC During Lunch. Middle Picture = Before; Right Picture = After

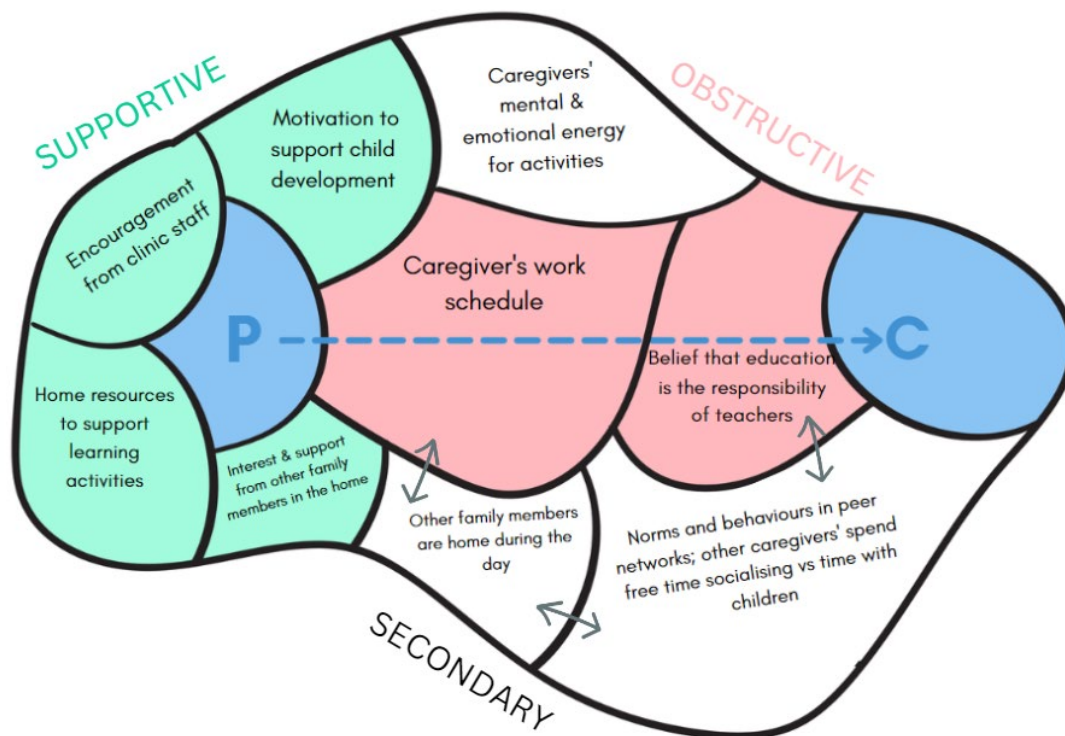


Context. As with the Scoping workshops, unpacking the programmes’ context was observed to be a rich and engaging discussion in both test cases’ workshops. This information came to the teams easily, which helped ground the ToC development in reality. The discussions were framed using life spaces, intended to identify the obstructive and supportive factors that could serve as realist contextual conditions in the ToC — i.e., the contextual conditions necessary to trigger mechanisms.

I leveraged my understanding of the ECD sector to prepare a draft life space model for discussion with the EduSpark team (see Figure 50). Heather collaborated with the EYEN team to create a map from scratch during the workshop. Despite the different approaches, discussions in both workshops were engaging and informative, although they were only loosely connected to the actual maps that had been created. The relationships between forces and the role of secondary forces were not the focus, as teams preferred to draw from their experiences to discuss context more broadly.

Figure 50

Example of a Life Space Shown to the EduSpark Team



In her final interview, Heather expressed an appreciation for this tool, as well as the process of unpacking the organisation's context, noting the importance of this exercise.

Indicating a preference for non-academic processes, she said about the life spaces:

"I thought that was a nice way to do that as well, which didn't necessarily need literature and all of that stuff. So, it doesn't make it necessarily academic ... and it was practical that they could do it themselves, based on their knowledge." (Heather, final interview)

Feedback. The final component of the SAMEL ToC is the incorporation of feedback loops, where teams are encouraged to identify non-linear relationships within their ToC (e.g., outcomes that reinforce or balance each other). This step draws on the first introduction of CLDs in the Scoping workshop.

Neither Heather nor I adequately attended to this concept in the workshops.

Heather's aforementioned discomfort with CLDs and feedback loops led her to only briefly mention the concept without delving into it, repeating her approach from the Scoping workshop — to "*plant a seed*" and swiftly move on without allowing the team to engage with non-linear relationships.

During the EduSpark workshop, time constraints prevented a detailed discussion on feedback loops. The team's in-depth engagement with defining their programme impact and interventions consumed more time than had been anticipated, necessitating an adjustment to the workshop schedule that limited the exploration of feedback. However, a brief mention of the concept provided some insight when discussing the causal links between outcomes and mechanisms. A team member recognised this and its relevance in understanding cyclical changes impacting the programme.

"I started making sense also of the concepts that Jessica introduced us to last week.

That one of the feedback. ... I could see how some of the things that are meant to bring change will actually maybe also create a cycle of feeding back to the programme." (EduSpark team member, ToC workshop)

Theory of Action Workshop

In the third workshop, the teams developed a ToA. As with the ToC, an entire day was afforded to the ToA workshop.

The SAMEL ToA consists of a participant definition, programme outputs, partners and their contribution to outputs and outcomes, external factors, and programme factors. As with the ToC, feedback loops are used to illustrate non-linear relationships between elements.

The systems thinking concepts proposed for ToA development are: a boundary critique of the programme's participant definition (inspired by CSH; Ulrich, 1983); actor maps

to identify the programme partners needed to achieve outputs (Gopalakrishnan & Clarke, 2015); life spaces to identify environmental factors that might impact programme implementation (Lewin, 1936); and service utilisation diagrams to identify internal/organisational factors that might impact programme implementation (Rossi et al., 2019).

Boundary critique. Both EYEN and EduSpark’s ToA workshop began with defining the active population of programme participants and critiquing the boundaries of the definition. In both workshops, I observed that comparing *ideal* and *actual* participant definitions stimulated powerful and sometimes divisive discussion.

Both EYEN and EduSpark had concrete eligibility criteria for their participant populations: ECD centre staff. Discussions in both test cases’ workshops revealed the implication of focusing solely on ECD centres rather than non-centre-based ELPs like community playgroups — the exclusion of children unable to afford ECD centre fees.

EduSpark’s CEO commented on how this session aided the identification of potential future strategic opportunities, saying:

“The aha moment for me actually was the, when we were doing the boundary critique. So, realising the potential that we have in that space, that could be what differentiates us, that could differentiate us from other SSOs.” (EduSpark CEO, ToA focus group)

In the EYEN workshop, one team member strongly emphasised the same point, arguing that provision needs to be made for children who cannot afford ECD centre fees and do not have access to subsidised community-based playgroups. Thus, these sessions surfaced an out-group — a subset of the child population that is not being adequately serviced by the organisations, and thus lacks access to quality early education. Staff were struck by the realisation, noting that their programme boundaries were creating an out-group of vulnerable

children. It is important to note the ethical implications of boundary critique, which has the potential to create tensions and raise emotions amongst team members. Heather later commented to me that the team member who made the comment appeared to be *“quite offended with the exclusion stuff”*. Despite this observation, in her final interview, Heather reflected on boundary critique as an important component of programme strategy, one that is often neglected.

Partners. Using a partner analysis template, the Toolkit encourages evaluators to help teams identify different types of partners (e.g., core, auxiliary, or collaborating), together with where the partners’ contributions fit into the ToA.

I observed that Heather's clear facilitation ensured smooth progress in the partner analysis session, with the EYEN team experiencing minimal confusion and high engagement during group work. After completing the partner analysis, the team identified the outcomes and outputs to which partners contributed in a plenary session. This exercise seemed to further reinforce the running joke amongst the EYEN team that they were *“extras”* in the ECD sector and not the *“main character”*. During the focus group, one EYEN team member remarked:

“Really sitting down and thinking, where do they fit? ... EYEN fits everywhere, we do everything — in our heads [laughter]. And now you break it down and you realise ... we are the secondary influences!” (EYEN team member, ToA focus group).

Heather’s approach contrasted with mine in the EduSpark workshop, where time constraints led to a more discussion-orientated session without completing the full partner analysis. Instead, the EduSpark team listed partner names per categories, and then collectively identified the ToA outputs to which each one contributed (which I marked with

sticky notes). The EduSpark team was observed to be attentive and interested in this discussion, which led to identifying opportunities for future relationships.

External factors. Life space models, introduced in the ToC workshop, assist teams to identify external factors that might impact their programme's implementation. These factors can either hinder or aid implementation, requiring mitigation or leveraging, respectively.

In the EduSpark workshop, time constraints led to a simplified approach to this activity. I was observed to list supportive and obstructive forces with the team, without exploring their interrelationships (a key aspect of life space models). In my journal, I noted that obstructive forces were identified more readily by the team, reflecting that it is *“easier for them to talk about because I guess it shifts responsibility away, and people like to vent about what makes their job hard”*. I also noted that this exercise felt repetitive, given its use in the ToC workshop.

The EYEN workshop also did not fully utilise the life space model's potential to illustrate the dynamics between different forces, as Heather did not explain the tool in great detail. Consequently, in groups, the team wrote linear lists of supportive and obstructive factors, and then copied them into a blank template without considering relationships between factors or the size of their influence (see Figure 51).

Figure 51

EYEN Team Member Draws a Life Space Model



While this exercise was not properly facilitated in either workshop, in my observation notes, I noted that having the EYEN team draw up the maps themselves (albeit incorrectly) was *“a lot more effective than looking at a pre-populated version”* (which was done in the EduSpark workshop), and helped to energise the team. During her final interview, Heather remarked that she particularly liked how the context and partners’ contributions were incorporated into the ToA.

In the EYEN focus group, the SEM (Bronfenbrenner, 1979) used in the Scoping workshop was noted as being useful for the life space activity:

“When we were busy with our enabling factors, we discussed crime as part of it, and poverty. But when I thought back to that, meso, what's it — ecosystem? I had to say to myself, 'No, Gail, you cannot solve crime in this country' [laughter]. You know? 'You single-handedly cannot stop poverty!'” (EYEN CEO, ToA focus group)

Programme factors. The SAMEL Toolkit recommends drawing service utilisation diagrams to identify internal barriers to programme implementation, which are then

translated into factors essential to enhancing service quality (that are controllable by the programme team).

In my journal, I noted some initial hesitation when I invited the team to direct their attention inwardly. The observer then recorded a *“moment of insight”* from one team member, who said, *“Our own quality and processes. I hadn’t thought of that.”* The team was observed to be attentive, and they engaged in *“self-reflection”*.

During the EYEN workshop, I observed that the introduction to this session was brief, focusing on *“internal factors”*, without a detailed explanation of the service utilisation diagram's structure. This caused some confusion. Despite this, group sessions fostered engagement and interest, although energy levels dropped in plenary discussions.

Feedback from the focus groups revealed an appreciation for this activity. The EduSpark team acknowledged that it helped them see how their own internal barriers may have limited their reach over time, and where there were risks that required mitigation through a *“relook at programme design”*. Similarly, the EYEN team valued the clarity it provided in identifying actionable intervention points, with one team member explaining:

“It gives you really key concrete places to intervene and fix what you're doing in a very straightforward way that, I think, a lot of people might not even consider because it always feels really overwhelming. When you look at it, like, 'Oh, it's communication. We can fix that'. I like that.” (EYEN team member, ToA focus group)

In her final interview, Heather also expressed interest in taking this concept forward in her consulting work, calling it *“beneficial and important”*.

Monitoring & Learning Workshop

“People leave, everything goes with them.” (EYEN managing director, focus group)

The objective of the Monitoring & Learning workshop is twofold: to devise a monitoring plan and to formulate a learning strategy. Both outputs are developed in the same day. These outputs should enable teams to collect data for adaptive programme management and create opportunities to reflect and use the data meaningfully. As illustrated in the extract above, institutional learning is not a taken-for-granted reality in organisations.

The Monitoring section of the SAMEL Monitoring & Learning chapter is traditional in the sense that it draws primarily on the work done in the Focusing workshops to develop a monitoring framework. As such, the systems thinking content in this section is more implicit than in previous chapters, as the focus is on measuring the systems-informed programme theory (e.g., monitoring contextual conditions and realist mechanisms). However, three additional 'light' concepts are introduced in the chapter, to broaden thinking around learning questions (using the systems thinking core concepts), indicators (using the CRISP acronym), outcome monitoring (drawing on participants' perspectives), and participant monitoring (distinguishing between those still involved in the programme and those who have disengaged).

The Learning section emphasises learning events — designated events where teams come together to reflect on programme- or sectoral data and formulate a way forward using the 3R Model and leverage points.

The Monitoring Plan

Both EYEN and EduSpark's Monitoring & Learning workshops commenced with identification of the areas of the programme theory (ToC and ToA) that would be regularly monitored, followed by developing learning questions to guide the monitoring plan. Despite being a recommended practice in MEL literature, the inclusion of learning questions in the Toolkit was a departure from Heather's and my usual consulting practices.

During the session on learning questions, both the EYEN and EduSpark teams showed signs of confusion and a noticeable drop in energy. To simplify the process, neither Heather nor I pressed the systems thinking core concepts approach to the questions, allowing the teams to focus on areas they deemed important. Heather's approach might have stemmed from her own challenges in grasping the framing of these questions using the core concepts, as mentioned to me in a pre-workshop check-in.

Subsequently, the teams moved on to developing indicators aligned with the programme theory. The SAMEL Toolkit advocates the use of both quantitative and qualitative indicators, moving beyond the traditional focus on quantitative metrics alone. It also introduces the CRISP acronym to guide the creation of indicators.

In practice, Heather primarily facilitated the development of quantitative indicators, while my efforts to apply the CRISP principles of sufficiency and pragmatism sparked debates. These discussions, particularly with two team members, centred on the practicality and feasibility of data collection in light of the organisation's capacity. I reflected on indicator development:

"Oh my lord, it takes forever. There's just so much. Every time I ask about the capacity to collect the data, Timothy says we should be planning for a best-case scenario."

(Monitoring & Learning journal)

During the indicator development session in both the EYEN and EduSpark workshops, questions surfaced on the specifics of the programme theory revealing underlying uncertainties. In the EduSpark context, I noted a persistent lack of consensus on the programme's objectives and design, highlighting foundational disagreements. Likewise, I observed the EYEN team raise issues with the outcomes that had been developed in the ToC workshop, with debates resuming around the issue of centre registration and conceptualising

quality improvement in their ToC. These discussions were influenced by the same contextual factors that had surfaced during ToC development, primarily ECD policy implementation. In my observation notes, I wondered: *“What does it mean that there is debate around the ToC at this point?”*

During a critical point in the workshop, the EYEN team encountered a significant misunderstanding regarding the structure of their programme. The programme manager introduced a perspective that suggested that what had previously been considered a single programme might actually be two distinct sub-programmes. This revelation led to considerable confusion. After a day filled with technical details and diminishing energy levels, it was as though a pressure valve had been released, with everyone’s frustration being released at once. In my notes, I somewhat dramatically observed:

“There is a collective meltdown. Zara walks out, comes back smelling like smoke. Kate says: ‘The next training we need is media training, because we cannot face the world like this. We are a mess!’” (EYEN workshop observation)

As the team collapsed into shouts of disbelief and shocked laughter, Heather gently admonished that further efforts were needed to define the programme's boundaries more clearly. When the managing director, who had been absent from this session, returned, the debate reignited as he was asked for clarity. I noted the following observation: *“Heather has kind of lost control of the room, without sounding dramatic. Sessions end dramatically.”* It appeared that the level of programmatic detail required by monitoring (e.g., specification of outputs) uncovered significant gaps in some team members’ understanding of the programme, and there was lack of agreement on the intervention details.

The Learning Strategy

In the EYEN session, the development of the learning strategy faced challenges due to the late start of the session and the aftermath of earlier fierce debates. The collective energy of the team was observably “sapped”, leading to a session in which the team struggled to maintain focus and coherence. I observed that Heather was rushed and somewhat frustrated.

“Session starts abruptly. Heather rushes through the introduction and reads off the slides really quickly. She mentions the learning strategy worksheet and says that she will leave it with them to look through... Discussion crumbles again. Heather reins in and continues reading off the screen. Doesn’t engage in any discussion. Doesn’t decide on learning rules”. (EYEN observation notes)

This was difficult for me to observe without feeling disappointment, particularly when the need for a learning culture surfaced several times in the focus group, for example:

“There needs to be a bit more process put in place, I think, you know. Yes, once that process is in place, and the data is there, it's more accessible, people know that they can use it, they know where to find it.” (EYEN managing director, focus group)

In the following Evaluation workshop, Heather briefly revisited the learning content that had been overlooked in the Monitoring & Learning workshop, dedicating about five minutes primarily to the 3R Model. I observed Heather’s delivery as monotonous, stemming from her reading directly off the slides, which lent a tedious air to the presentation. This contrasted with her sentiments surrounding this section. In her final interview, she reflected:

“The 3 Rs and building in learning events, I really like that. And that is something that I want to take forward in my practice. Because that's also new, like building in a specific learning plan. I really, really like that.” (Heather, final interview)

Conversely, EduSpark's session on this topic unfolded more smoothly, and was met with enthusiasm and high levels of engagement. In my journal, I wondered whether this was because the session was *"not so technical, but more strategic and common sense"*, compared to the prior sessions. Discussions with the team underscored the value of cultivating a learning organisation, a quality sought after by donors and historically undervalued within the organisation itself. Reflecting on the need for a learning culture, EduSpark's CEO commented:

"It should be everybody's responsibility. I believe in that strongly. So, that is, to me, just something that I liked out of the session, just reminding us that there is no learning culture in this organisation." (EduSpark CEO, focus group)

The EduSpark team showed interest in and an eagerness to use the 3R Model and learning events, as well as the SWOT analysis (which they noted was useful for reviewing partners' memoranda of understanding). However, the concept of leverage points did not appear to resonate as strongly with the team. Heather did not introduce leverage points to the EYEN team, but, despite my introduction being brief, there was some appreciation for the 3R Model in the EYEN Evaluation focus group:

"It's like almost thinking about what you're thinking about. Thinking about what you're doing. And understanding that Triple R, how incredibly important that is, because, otherwise, we're just doing the same thing all the time." (EYEN CEO, focus group)

Evaluation Workshop

"[Evaluation is] above the sky, it's the — like the sun, you know ... It's one thing that you cannot reach ... So, for me, I think it broke that stigma or that assumption of like, 'Yoh, it's scary up there'." (EduSpark team member)

This captures the common perception amongst NPO teams that programme evaluation can be daunting or even untouchable. The SAMEL Evaluation workshop aims to

demystify this process and prepare teams for hybrid evaluations combining internal and external expertise. It is aimed at instilling best practices and confidence in participating in the evaluation process.

The workshop has two main objectives: to develop an ideal evaluation plan and to enhance team readiness for evaluation. The plan encompasses objectives, criteria, questions, and making 'ideal' boundary decisions using CSH (Ulrich, 1983). This approach helps compare the ideal with the actual evaluation plan by addressing 12 key questions .

To boost readiness, the workshop introduces fundamental Made in Africa principles. For this study, a board game was designed for this part of the test. This game acquaints teams with hypothetical evaluation scenarios, intended to make the evaluation process more accessible and less intimidating.

The Evaluation Plan

Both workshops commenced with an introduction to the MAE principles (Chilisa & Mertens, 2021), which immediately captured both teams' attention due to their contextual relevance. There was an appreciation for MAE in both focus groups, with the teams indicating a desire for solutions grounded in local realities. For example:

"I fell in love with MAE — Made in Africa Evaluation — because I always felt like some of the things that we try to apply in our communities are not relevant context-wise. So, I really like the fact that we are beginning to look at what is happening here and trying to make a difference with our own solutions." (EYEN team member, focus group).

Interestingly, Heather's impression of MAE contrasted with that of the teams. In her final interview, she said that MAE was *"like a really academic input first thing that probably switched people off"*.

Following the introduction to MAE, both workshops moved on to criteria-setting and definitions.

In the EYEN workshop, Heather appeared to show a preference for the traditional DAC (OECD, 2021) criteria. I observed her focus on these more familiar criteria as she described the SAMEL criteria as “*more focused on the South African context*” (but not explaining further).

Both teams were observed to grapple with the criteria, as competing ECD priorities appeared to cause confusion. For example, the registration debate that characterised the EYEN ToC workshop re-surfaced as the team debated the role of registration in the criteria definitions. Similarly, during the EduSpark workshop, the Connection criterion raised a debate around advocacy and whether it should form a part of EduSpark’s interventions.

Criteria-setting also appeared to stoke tension between ideal reporting standards and the realistic contextual constraints of NPOs. For instance, the EduSpark team was observed heatedly debating the target-setting component of criteria definitions. Disagreements surfaced over what constitutes an achievable success target, with some suggesting the proposed targets were too ambitious for their context.

In the EYEN workshop, discussions focused on the scope and specifics of criteria. I recorded that the CEO was “*very preoccupied with proving as big an impact as possible*”, as the CEO referred back to the SEM’s (Bronfenbrenner, 1979) macrosystem. This seemed to strike a chord with the MEL manager, whom I observed lamenting the difficulty of measuring criteria.

Following criteria definition, the teams posed evaluation questions. As with the Monitoring & Learning workshop, I recorded some confusion amongst the EduSpark team, while the observer recorded “*a sense of quiet*”. I was observed prompting the team using the

example questions provided in the Toolkit (using the systems thinking core concepts), which Heather also provided to the EYEN team. However, neither team incorporated the core concepts into their evaluation questions, choosing to focus more on content areas related to their programme theories.

The final session to complete the evaluation plan involved discussing the CSH system-bounding questions in the ideal (*ought*) mode (Ulrich, 1983). This activity is meant to help teams plan the ideal circumstances of an evaluation, which they can later compare to what is actually feasible, including aspects such as expertise, decision making, and meta-evaluation.

In my journal, I described the session as smooth, with high engagement and understanding amongst the EduSpark team. Discussions about potentially negatively impacted stakeholders in an evaluation were noted as particularly insightful, as was the value of diverse perspectives in a steering committee. The team seemed to grasp CSH's principles naturally. A standout moment was when an EduSpark team member suggested that the evaluation report include the views of those who disagree with the evaluation findings, embodying CSH values. I reflected on this moment in my journal:

"I LOVE THAT IDEA!! Such a goodie, and a really nice application of CSH in its own way — opening up that boundary and very deliberately including perspectives — opposing ones, no less." (EduSpark Evaluation journal)

In the EYEN workshop, I observed that the team answered the CSH questions "*fairly easily*" and with a good comprehension of the questions. However, Heather's explanation of the *ought* vs *is*' distinction and its relevance to evaluation planning was unclear. The EYEN team appeared to find the exercise's idealistic approach challenging, as the managing director frequently mentioned the contextual difficulties the evaluation would encounter, emphasising the need to be less idealistic and more realistic. My notes captured a moment:

“Mandla says, with a cynical smile: ‘Ideal situations make life so much easier.’” (EYEN observation notes)

Despite there being time to finish answering the questions, Heather ended the session abruptly at the tenth question, saying: *“I can feel the energy in the room is very low.”*

Team Readiness

After developing an ideal evaluation plan, both teams engaged in a board game designed to highlight common evaluation challenges and promote collective decision making. Both teams received this game with enthusiasm.

In the focus groups, team members from both EYEN and EduSpark praised the game as both enjoyable and instructive, and noted that it facilitated group discussions and prepared them for potential future challenges. An EduSpark participant highlighted the game's effectiveness in reinforcing the Evaluation workshop's content, while EYEN members appreciated the emphasis on shared problem-solving and finding “common ground”. For example:

“I liked the fact that you couldn't move on until you agreed. So, I think that was really important. You get to share your opinion. But, again, it pulls you back to everyone being on the same page before you can move ahead.” (EYEN team member, focus group)

In the EYEN focus group, a team member appreciated the game's role in preparing them for evaluation realities. They emphasised the need for planning tools suited to the specific fieldwork context, echoing earlier discussions on MAE.

The game positively impacted the atmosphere in both sessions, infusing teams with energy, humour, playfulness, and lively discussion (see Figure 52). This appeared to be an

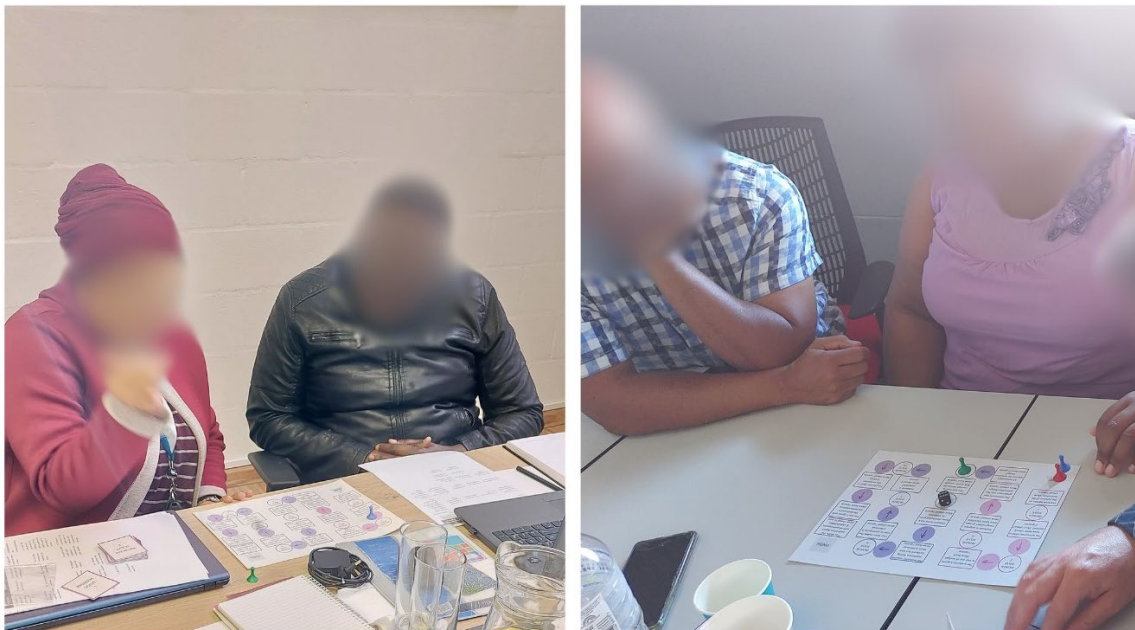
important value-add to a day that was “heavy” on technical content, as explained by an EduSpark team member in the focus group:

“I quite like the introduction of a game because it kind of stimulates you, brings you out of the heaviness a bit ... I needed the game to get me out of the heaviness.”

(EduSpark team member, focus group)

Figure 52

EYEN Team Members (Left) and EduSpark Team Members (Right) Play the Board Game



After the game, the workshops concluded with a debriefing session. Heather and I guided discussions on evaluation readiness using a worksheet with prompts regarding the programme stage, data- and resource availability, and team expectations. In my journal, I noted that the EduSpark team's discussion was somewhat limited, possibly due to fatigue or, I theorised, “because evaluation feels very far away”. There was confusion amongst the EYEN team about how to approach the exercise. I noted one member asking, “How are we supposed to answer the questions?”, as they did not have the necessary data. Little

explanation was given by Heather as to the purpose of this reflection and when it should be repeated (prior to a planned evaluation). I observed that the team was both distracted and full of energy following the board game, possibly contributing to the low engagement observed in this session.

Heather's impression of the Evaluation workshop as a whole was that it took place too soon in the MEL development process. I noted that, in the final framework she had sent to the EYEN team, she did not include any of the work done on the day (such as the team's answers to the CSH questions), choosing instead to provide excerpts from the Toolkit to guide the team when they plan an evaluation later on. She said that this was because she felt the team was not ready for evaluation preparation, and that *"the workshop gives them the tools to develop it later"*.

Workshop Aftermath

After concluding the EYEN workshops, I conducted a debrief interview with Heather to reflect on the Toolkit as a whole and the workshops she had facilitated. Two months after the completion of the SAMEL workshops, I conducted follow-up interviews with EduSpark's CEO, EYEN's CEO, and the MEL manager to gather information on what had transpired after the SAMEL workshops.

Heather's Reflections

Heather felt that the Toolkit successfully instilled an understanding of systems thinking within her, as well as its relevance to MEL, saying that it broadened her perspective, helped combat siloed thinking, and helped her appreciate the relevance of the systems in which a programme operates.

The systems thinking tools she found particularly helpful and interesting were the SEM (Bronfenbrenner, 1979), boundary critique inspired by CSH (Ulrich, 1983), life space models

(Lewin, 1936), incorporating partner inputs and contextual factors into the ToA, the service utilisation diagram (Rossi et al., 2019), the 3R Model, and learning events.

The presentation and format of the Toolkit seemed to be approachable and user-friendly, as Heather remarked:

“I think how you've done it graphically is really nice. ... like the extra bits of input that you've put in in different formats. I think the language that you use is good. It's not all academic; there's parts that make you smile.”

The extent to which she is likely to significantly amend her professional practice based on what she learned is, however, questionable. Heather remarked that she would like to emulate SAMEL's approach to evaluation preparation and readiness in her MEL consulting, as well as the programme strategy foundation (problem statement, vision, mission, and values). Heather also mentioned taking the service utilisation diagram and 3R Model forward in her work. None of these tools are systems thinking tools from any particular methodology, but they are complementary to MEL. One possible reason for this hesitance in taking systems thinking forward in her practice is the perception of systems thinking or the Toolkit process as incongruent with her preferred way of working.

Heather regularly referred to the Toolkit content as “academic” throughout the workshop. In her final interview, she reflected on the Toolkit:

“It was useful. But it was academic, for me. This, like, the whole process is more academic than I'm used to at this point. And that's why I think I probably attacked the chapters in the way that I did, because I want to, like, do things, I want to action it, I want to you know, like, get the job done, like I do with my clients ... So, yeah, ... a lot of it was a lot more academic than I'm used to. And I think that's also part of the process

of the systems thinking, ... so many other things to take into account than your narrow way of working, which I currently have, so that was a good thing."

In the extract above, there appears to be a tension between Heather's view of the utility of systems thinking as broadening one's perspective and the practical downside of this learning. A broad perspective takes too long to practically action with her MEL clients, or does not fit into her preferred MEL consulting routine. This was alluded to in the extract below, in which Heather described how difficult it is to change one's routine:

"I already have my way of working, and I've been doing it for X amount of years ... We had this discussion around it being harder to change your practices when you're so far in ... which is why I probably wouldn't incorporate all of the systems thinking, but it definitely raised my awareness around it and appreciation of it."

Heather's approach is perhaps representative of the tension between learning a new approach versus resisting professional change. Her engagement with the SAMEL chapters was described in the interview as Heather's pre-workshop preparatory method. She described this method as follows: (a) skim-read the first sections, paying attention to headings; (b) created a slide deck structure using the evaluator preparation guidelines; (c) went back to the beginning of the chapter to read the content, pasting relevant information into her slides; and (d) completed the worksheets. Heather stated that this approach made her feel more comfortable, as she knew what she was going to facilitate in the workshops. However, it also meant that I observed Heather reading off her slides in every workshop, demonstrating, at times, poor relational understanding of the content.

Given Heather's approach to workshop preparation, her ability to fully internalise the Toolkit content is brought into question. It does not seem as though the content was read for

understanding, but rather for sharing in the workshops. This appears to be supported by the following extract, which indicates that she seemingly felt lost at the end of each chapter:

“Without the worksheets, I wouldn't have known where to start. Well, not wouldn't have known, but it would have been harder. So, the worksheets were helpful. And I think the more practical ones that I could do in the workshops were really helpful, because it laid out what you needed to cover.”

At the end of this extract, we can again see Heather prioritising workshop facilitation over her own internalisation of the content.

Thus, while Heather appeared to see the value in systems thinking, there appeared to be a reluctance to engage with the content meaningfully and for a period sufficient to ensure that she internalised it. The priority was preparing for the EYEN workshops. A strong possible explanation for this is simply time.

Heather's engagement with the Toolkit was in the context of the test, as opposed to a completely voluntary and self-directed professional development exercise. She needed to engage with the SAMEL Toolkit, plan the EYEN workshops, and then facilitate the workshops within a tight timeline (four months), all while conducting other consulting projects. On this workload, she commented:

“I think that's what made it difficult. It's just that I was trying to learn new things around juggling a million other clients and wanting to do it justice. So, I think that was the struggle, more than the content.”

Heather's workload is typical of MEL consulting, in my experience at least. It can be difficult to prioritise time for professional development and upskilling. Heather and I discussed this:

Heather: *"In general, I don't make the time for upskilling academically, but I want to, which is why, when this presented itself, I said yes."*

Me: *"Okay. So, would you say time is your biggest barrier?"*

Heather: *"Yes. One hundred percent".*

EduSpark

Following the workshops, the EduSpark CEO, Dion, noted a *"sense of relief and accomplishment"*, and described the engagement as *"intense"*. Dion reflected on the major takeaway from the workshops: a greater depth of understanding of the organisation's implementation context, and feeling better positioned to offer the team strategic support:

"The process with you allowed me to understand the nitty gritty of the implementation of the programme much better. So now, when Zola speaks about stuff, I understand, and I can connect with the things much clearer and better, and give more strategic support and advice now."

Dion described the team feeling positive and excited about the ECD programme model they had designed in the workshops, and noted that they were making an effort to share it across the organisation, as well as with the ECD principals who had participated in the Scoping workshop. The latter was deemed an important step for EduSpark, as their ECD programme had, until recently, been unstable, and was facing closure. It was important to the team to allay fears amongst their stakeholders.

Dion described how the SAMEL process came at the right time for the organisation, as it was in a transitional period. The workshops helped support the strategic revisioning they needed to do. Dion praised the *"structured"* manner in which the MEL framework was developed throughout the SAMEL process, appreciating that *"it holds people accountable"*.

Dion noted that this was particularly important in a difficult and unstable organisational context:

“From an organisational perspective, it's been a very enriching, but also a needed intervention for the organisation's survival and growth going forward, you know, on various levels. I mean, we want to attract the big funders, and I'm sure that's going to come”.

With regard to the utilisation of the final framework, the focus appeared to be on the programme's design. Dion described how the ECD programme manager immediately incorporated the programme theory in staff training. Dion also described using the extracts from the framework, particularly the vision, mission, and ToC, to change the organisation's constitution, funding applications, and other strategic engagements with stakeholders. The process as a whole appeared to have given the organisation legitimacy or a sense of professional credibility, as Dion illustrated through an exchange with a potential funder:

“He was quite impressed with the fact that we have done this process with you.”

Dion shared a number of barriers to utilisation, most notably in terms of organisational capacity and resources. Dion lamented not having a formal MEL manager to help drive the implementation of the MEL framework, and noting that the learning strategy had been neglected following the workshops, as the learning champion that the team had nominated had left the organisation. The need for MEL support was focused on reporting, not learning. Dion said:

“I mean, now, just for the annual report, we just realised, once again, the importance of having your monitoring and evaluation framework in place.”

Finally, Dion expressed regret about not including more staff in the SAMEL workshops to get greater buy-in and understanding. He noted the difficulties experienced by staff who

were not “*part of the dynamics*” in fully understanding the MEL framework, noting that, while the final document was shared with the team, they might not have read it in full.

EYEN

The immediate priority following the SAMEL workshops for the EYEN team was setting programme deliverables that need to be managed in EYEN’s partnership with the DBE. Zara, EYEN’s MEL manager, said that a logframe was created after the workshops to share these deliverables and time frames with the DBE. The CEO, Gail, went on to say that the team was “*still busy with the programme design*” using the ToC. They reflected on the strategic clarity that they got from the ToC development, noting that they normally designed programmes “*back to front*”, beginning with the design and then later deciding on what outcomes they wanted to achieve based on the design.

Regarding utilisation, Gail said that this front-to-back design logic had been applied to the other programmes in EYEN’s portfolio, to create greater strategic alignment across all of the programmes. Gail and EYEN’s managing director had been working on the organisation’s overall strategy, with which all the programmes would need to align. Gail said the ToC designed during the SAMEL workshop would probably be revised once this organisation-level work had been done, noting a need for a more cohesive set of organisational objectives across programmes. The team decided to continue their work with Heather, employing her as a consultant to help them repeat a similar process for their other programmes. When I probed further, Zara said that the process would be simpler than the SAMEL test that was conducted for this study, as the SAMEL workshops helped prepare the team to be more “*directed*” in terms of the overall strategy.

While looking at the design of another of EYEN’s programmes, Gail described how the team reflected on the situation analysis, particularly the SEM (Bronfenbrenner, 1979). Gail

said they were better able to focus their attention and identify areas of influence versus what should be left to partners. Zara went on to describe how the team was looking to strengthen programme partnerships following the workshops, with Gail adding that the team had a greater understanding of the role they play and the value of collaboration. For example:

“It's reinforced that there's no competition necessary, because none of us can actually tackle that — anything on our own, and that, when we work together and work in a collaborative way ... then we can address the whole ... Your little role in it maybe seems small, but when you put it all together, it works very cohesively.”

Gail and Zara reported that the programme theory developed during the two Focusing workshops had been used as part of funding proposals, as Dion had noted doing with the EduSpark strategic work.

As with EduSpark, Zara noted some challenges with regard to implementing the MEL framework, saying that this had not yet happened in any meaningful way, due to scheduling difficulties with the greater team and *“lots of moving parts”*, particularly with the DBE who wants data collection to take place digitally. Zara noted that the team was *“just trying to stay up to date and align with what's happening in the moving parts that we don't have control over”*. Echoing Dion's difficulties with sharing the MEL framework with the broader team, Zara indicated only sharing snippets of the MEL framework with certain team members, so as not to *“inundate”* them with too much information.

To improve MEL implementation, Gail noted a pressing need to get organisation-wide buy-in, saying:

“I shouldn't even call it M&E, we should call it the continuous, you know, learning, because, sometimes, I just leave it in the hands of one or two people ... We need to find a way where people realise that everybody's responsible for this.”

Further reflecting on the need to institutionalise learning, which did not result from the SAMEL process, Gail noted a pressing need to gain more operational support, because the team was *“so thinly spread”*, adding:

“Now that we're trying to extend, expand our capacity, we can afford to become more focused ... Because you can't just go and do, do, do, and then come up for air, and then figure out 'Actually, are we going in the right direction?’”

Key Feasibility Issues

This section summarises the testing of the SAMEL Toolkit, highlighting the interplay between conceptual understanding, environmental and contextual factors, interpersonal relationships, and logistical challenges. Several barriers and enablers are noted that speak to the feasibility of implementing the SAMEL Toolkit in real-world settings, underscoring the importance of several aspects of facilitation.

Facilitator application of concepts

A significant issue that emerged in both tests was the challenges that Heather and I faced in practically applying the SAMEL content in workshop facilitation. We both particularly struggled with causal loops and integrating system feedback into the workshop sessions, demonstrating low self-efficacy and confidence in our ability to explain the concepts to the teams clearly. To avoid confusion and time delays, we both pivoted the sessions when we did not feel confident, reverting to more linear models of explanation and discussion, or abandoning systems thinking concepts altogether. Challenges with practical application suggest that the Toolkit may have successfully transferred conceptual knowledge, but not procedural knowledge — knowing *how* (Clinton & Hattie, 2021).

Atmosphere

The energy and atmosphere in both workshop environments fluctuated significantly, which influenced team engagement. Some contributing factors were uncontrollable, such as hot weather causing drowsiness and workshops taking place during Ramadan. Other factors were more controllable, such as taking comfort breaks and using icebreakers or check-ins at the start of workshops. The Toolkit content also influenced the atmosphere. Concepts that allowed for group work or direct interaction (e.g., life space modelling) energised the teams, while technical and conceptually challenging content sessions tended to lower their energy and engagement. This supports the view that monitoring and managing a group's energy using creative techniques is a core facilitation skill in enabling quality participant inputs (Mackewn, 2008).

Rapport and Facilitation Style

Heather and I both developed a strong rapport with the teams. We were both observed to facilitate with respect and inclusivity, but also assertive probing, which the teams appreciated, as it enhanced clarity and stimulated critical thinking. Strong facilitation is characterised by the ability to apply polarised skills in practice, such as knowing when to listen and probe versus when to direct (Mackewn, 2008). While there were difficult moments, the atmosphere of the workshop sessions overall was good-humoured and characterised by trust and patience, allowing for constructive and respectful discussion and debate.

Team dynamics

Several workshop sessions, especially the more technical sessions (e.g., the monitoring plan), were dominated by particular team members. At EYEN, the CEO and MEL manager tended to speak the most, with significant input from the managing director. At EduSpark, the same was observed with the ECD programme manager and fundraiser/interim

MEL manager. This dominance sometimes led to narrowed discussions that may not have fully represented the diversity of perspectives within the team, which is a common issue in the workshop format (Maini et al., 2018). However, due to time constraints, Heather and I both allowed this to occur in more technical sessions, to be able to extract the information we needed as quickly as possible. However, this led to adequate involvement of all team members.

Time constraints

Time constraints notably impacted the ability to cover all workshop content thoroughly. These constraints were often due to late starts to the day and extended team discussions and debates. In many cases, Heather and I allowed discussions to run over their allocated time, as it was clear that these conversations were critical for the teams. This is a recommended practice in workshop facilitation, but it necessitates a degree of adaptation and flexibility to suit participants' needs (Goremucheche, 2022; Mackewn, 2008). In addition to workshop time, the time frame allotted to the test was a notable cost to the teams, who had exceedingly demanding operational schedules. While intended to be a beneficial strategic process, the development of MEL frameworks can be a heavy burden for non-profit teams (Adams & Garbutt, 2008).

Chapter Conclusion

This chapter provided an in-depth exploration of the testing of the SAMEL Toolkit using the EYEN and EduSpark teams, drawing on multiple data sources to present a comprehensive case description.

The feasibility of a systems thinking approach for MEL development in NPO contexts was demonstrated through the test, albeit with significant challenges. Both NPO teams found the SAMEL Toolkit to be strategically valuable, with the Scoping and Focusing workshops

providing particularly useful insights into programme design. The Evaluation and Monitoring & Learning workshops seemingly bolstering the teams' MEL competencies and sense of empowerment, and motivated them to own the MEL role. However, the theoretical complexity of some systems thinking concepts, particularly non-linear relationships, posed difficulties in the workshop facilitation. Furthermore, time was an ongoing issue before and during the workshops' implementation. Persistent time constraints affected the depth of content coverage, necessitating flexibility and adaptation in facilitation to accommodate critical team discussions.

The next chapter reports the findings of the reflexive thematic analysis of the test data through an interpretive lens, providing a richer understanding of the factors affecting the integration of systems thinking into MEL development in the NPO context.

Chapter Nine: SAMEL Test: Thematic Analysis Findings

The previous chapter provided a narrative account of the testing of the SAMEL toolkit through a description of what happened in each workshop and the factors that affected the process. This chapter takes a more interpretive approach to the test data by reporting on the reflexive thematic analysis of the same data (Braun & Clarke, 2006, 2022). This was done to understand the NPO teams' experience of the Toolkit content and to answer the following research questions: (1) How did the workshop participants and the evaluators (workshop facilitators) experience the SAMEL Toolkit as a systems thinking approach to MEL? and (2) What are the challenges in and advantages of integrating systems thinking into MEL?

Thematic Overview

Three themes emerged from the reflexive thematic analysis of the test data. I begin this chapter with high-level descriptions of each theme and the relationships between them, to prime the reader for the sections that follow. Please refer to Appendices L and M for a summary of the thematic analysis, including code descriptions.

Theme 1: Strategic clarity (Programme Theory)

This theme captures the primary advantage of the SAMEL process: an opportunity for teams to come together to engage strategically, in a structured manner, and develop a programme theory enhanced by systems thinking. The theme illustrates the significant value that the first three workshops created for the NPO teams — the development of programme theory as a conceptual tool for responsive programme design and strategic planning.

Theme 2: Contextual disconnect

This theme encapsulates the ongoing tension between an “academic process” (i.e. theoretical) and its practical application in settings with limited resources. Non-profit work involves providing much-needed services with limited resources in complex social contexts,

and is characterised by interactions between environmental constraints, policy pressure, and donor influence. This theme explores the sense of overwhelm amongst NPO teams and the implications of complexity for the development of programme theory. The team debates and discussions around programme design were lengthy, and issues remained unresolved across the workshops, constraining MEL development based on the programme theory.

Theme 3: Organisational disconnect

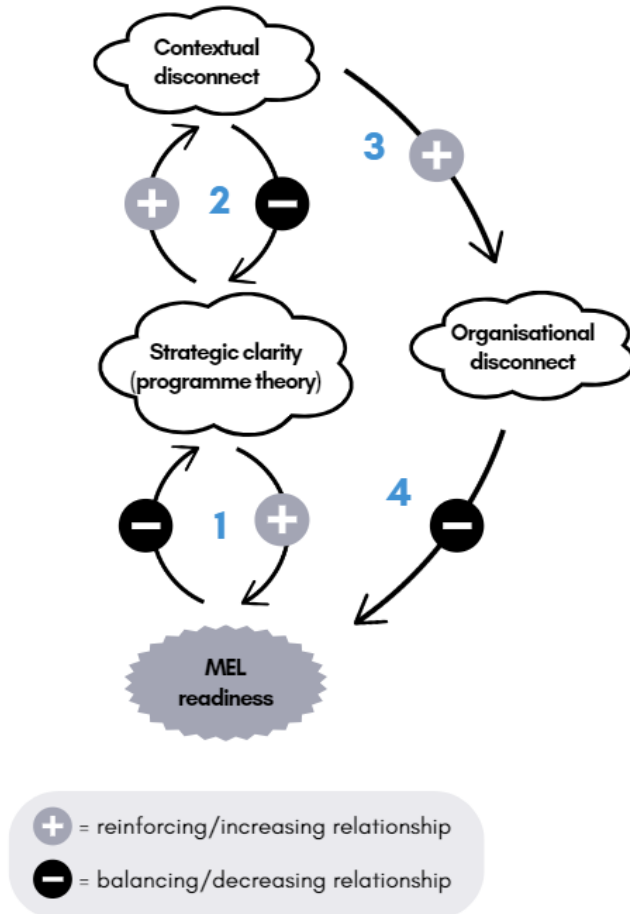
This theme highlights the organisational features that emerged from complex sectoral contexts that may constrain MEL processes. It explicitly illustrates the organisational barriers and enablers of MEL readiness (observed in the tests) that affected the SAMEL process.

Relationships between Themes

While these themes are presented in a linear sequence, it is important to note their non-linear and reciprocal nature. As an extension of the systems thinking lens applied to this thesis, I illustrate these relationships in the causal loop diagram in Figure 53, which highlights the reinforcing and balancing feedback relationships between the themes. These relationships are summarised below and elaborated upon in the remainder of the chapter.

Figure 53

Four Non-linear Relationships between the Three Themes



In the figure above, the three themes are contained in the clouds. A plus (+) sign indicates a reinforcing relationship between themes, while a minus (-) sign indicates a balancing relationship. The numbers in Figure 53 indicate four relationships, described below using the corresponding numbers.

- 1 Strategic clarity, through a systems thinking-informed programme theory (Theme 1), enhanced some aspects of the teams' MEL readiness. However, other aspects of MEL readiness that were not present in the organisations constrained the development of the programme theory (e.g., organisational capacity and buy-in for MEL).

- 2 Strategic clarity enhanced the teams' ability to deal with the contextual complexity of their environments (Theme 2), but this complexity affected the teams' ability to resolve the programme theory in the allotted workshop time.
- 3 The contextual disconnect between the NPOs' complex contexts and the SAMEL process led to Theme 3: Organisational disconnect. The resource-deprived nature of NPOs means that these organisations lack capacity or resources for comprehensive MEL.
- 4 Consequently, organisational factors (which resort under Theme 3) constrain teams' ability to be MEL-ready.

Discussion of Findings

The three themes presented above are discussed in detail in the following sections, supported by quotations from the workshop data (in addition to those already provided in previous discussions).

Theme 1: Strategic clarity (Programme Theory)

The SAMEL process had a notable advantage that was highlighted in both tests: the workshops fostered powerful engagement with strategic programme design and planning. In particular, the Scoping and Focusing workshops appeared to be the most impactful, highlighting the value of systems thinking concepts in strategic planning (particularly regarding situation analysis and programme theory).

Regarding the Scoping workshops, the EYEN test data indicated that the SEM (Bronfenbrenner, 1979), CATWOE (Checkland, 1981), and actor mapping (Gopalakrishnan & Clarke, 2015) were powerful reflective tools, especially when applied in combination. The SEM helped the team appreciate the size and scale of the ECD sector and where they fit into the broader context, while the actor mapping and CATWOE sessions helped them understand

what is within their reach and what role they can play effectively in the sector in relation to other stakeholders. In essence, the NPO teams felt a sense of clarity regarding strategic boundaries — what the organisation can and should do, and what other actors might do better. In the extract below, this is alluded to by EYEN’s CEO, who described the need to “let go” of some aspects in order to focus more strategically on their role and areas of reach:

“Just to say that it's created clarity. But I think clarity more ... in terms of what the focus needs to be, and the approach. And sometimes it's hard to let go. But it's also made me realise that, actually, you need to. It's necessary.” (EYEN CEO, Scoping focus group)

The EduSpark test also demonstrated the power of in-depth engagement with the programme context. This was particularly evident in the team’s understanding of the challenges faced by the local ECD community, a perspective enriched by including ELP principals in the Scoping workshop. Discussing problems and their causes in the CLD session (after the notes on the whiteboard had been linked together) was a powerful way for the team to obtain rich insight into the community's needs, instead of relying on their assumptions. Sentiments in the focus group discussion centred on these on-the-ground insights and the need to weave them into a responsive programme design (supporting quotations were provided in Chapter Eight).

For both teams, the Scoping workshops appeared to provide grounded insights into focus areas for their programmes’ design. These insights then fed into the Focusing workshops on programme theory, in which ToC and ToA development helped them hone in on these focus areas.

The focus group discussions following both test cases’ ToC workshop centred on the power of bounding and defining the interventions and outcomes, and the teams deliberately

engaged on the goals of their programmes. Fleshing out mechanisms and outcomes helped the teams collectively decide on a focused set of priorities for their programmes. The EYEN CEO remarked:

“It focuses us more on ‘these are the outcomes’. So, you’re not going to do or engage in activities that don’t meet those outcomes.” (EYEN CEO, Scoping focus group)

The ToC workshop also provided a critical opportunity for the respective teams to come together to deliberate on their programme designs and identify gaps in the programme models. For example, in the extract below, EYEN’s programme manager illustrates the relationship between outcome identification and intervention design:

“It’s more clear to me — ‘What goals do I have?’ What goals I need to achieve in my programme. And where my gaps are.” (EYEN programme manager, ToC focus group)

The opportunity to collectively discuss programme design and goals was a critical step for the EduSpark team, as they had never had such an opportunity. The CEO said:

“We’re sitting with information; we’ve got it written out. Bits there, bits there, some in more detail than others, and a lot in our heads. And we haven’t gone through a structured process of unpacking all of that and processing all of that information — ideas, thoughts, vision — into such a structured programme”. (EduSpark CEO, ToC focus group)

The CEO went on to describe how this process had helped to facilitate the team's ownership of the programme design, a sentiment echoed by the EYEN managing director, who emphasised the importance of team members contributing to model development and having *“a piece of skin in the game”*.

The second Focusing workshop (on ToA development) helped enhanced both teams' clarity regarding strategy, and provided additional opportunities to critically appraise their

programme designs and boundaries. The boundary critique sessions, in particular, raised ethical implications regarding the programmes' target populations, which led to powerful debates amongst both teams. The ToA workshop also enabled the teams to strategically consider what partnerships were important for service delivery, drawing on the actor maps completed in the Scoping workshop. This appeared to help the teams manage the complexity of their work, as noted by the EYEN CEO:

“So, once we honed in, it helped us to structure our thinking a lot more on the critical and what's necessary, and not feel overwhelmed by the stuff we're not doing.” (EYEN CEO, ToA focus group)

Finally, drawing service utilisation diagrams was a powerful exercise for both teams. They looked internally and identified specific programme implementation barriers emanating from the organisation itself. Both teams reflected on the notion that implementation barriers may be something controllable — or even caused — by the teams themselves, referred to as *“blind spots”* by an EduSpark team member:

“She flipped the coin and said, Okay, what then? What now? We didn't think of our own blind spots!” (EduSpark team member, ToA focus group)

Team members reflected on the importance of these barriers, with one EduSpark team member saying that losing participants due to these issues was an *“intense”* loss, as the programme might have significantly improved the life of the lost participant. Not only did this exercise surface these issues, but it appeared to make their mitigation more manageable by identifying *“straightforward and concrete places to intervene”*, in the words of an EYEN team member.

Lastly, the Scoping and Focusing workshops were impactful in creating a shared understanding of the programmes and their contexts. This was especially important because

both teams had recently recruited new managing members who needed a detailed understanding of what the programmes did, and why. This is illustrated by the extracts below. The EYEN MEL manager and EduSpark CEO, in reflecting on the SAMEL process during their follow-up interviews, said:

“I do think that, across the board, and again, because we have lots of new people coming into the organisation, but also there's been changes in the programme over the years. And so it really, from my point of view, I think it really brought everybody onto the same page. It allowed us to all have a similar understanding of what the programme is meant to do and how we are envisioning it to reach the outcomes we want it to reach.” (EYEN MEL manager, follow-up interview)

“The process with you allowed me to understand the nitty gritty of the implementation of the programme much better. So now, when Zola speaks about stuff, I understand, and I can connect with the things much clearer and better, and give more strategic support and advice now, because I have a better understanding of it [the programme].” (EduSpark CEO, follow-up interview)

In summary, both test teams saw immense value in the Scoping and Focusing workshops, particularly with regard to gaining clarity on the programmes' envisioned impact and outcomes, as well as the most feasible and appropriate boundaries for the programmes, based on their environmental contexts. This then helped teams to identify future opportunities (including stakeholder relationships), as well as potential implementation barriers around which they needed to plan. Discussions in these workshops were rich, animated, and nuanced, and they appeared to bring to the fore critical programme design decisions that had not been attended to up until that point.

Theme 2: Contextual disconnect

Perhaps an inescapable reality of MEL development processes in non-profit environments is the influence of their environment on the process. Data from the SAMEL tests captured the complexity of non-profit work as the NPO teams grapple with MEL development against a backdrop of extreme poverty, an immense need for community services, policy demands, and donor pressure. While the Toolkit enabled productive engagement with this complexity, I, in designing the Toolkit, had underestimated the consequences of this engagement with complexity in moving through the MEL development process as planned.

Surfacing and grappling with contextual complexity can be conceptually overwhelming for NPO teams accountable for service delivery under these conditions. In the extract below, EYEN's CEO expresses how overwhelming non-profit work can be in the face of urgent community needs and resource constraints.

"It's extremely daunting. It's extremely overwhelming. It almost feels like, you know, 'I don't think it's gonna work. Let's just throw in the towel. Let's just run away'. Because that's quite often how you get to feel. It's too much." (EYEN CEO, Evaluation focus group)

The word *overwhelm* was used several times by various test participants (and by me), during the workshops and afterwards, as we all tried to conceptualise the programme and MEL design in such challenging contexts. EduSpark's CEO described the team feeling overwhelmed after the workshops as they faced implementing a new ECD programme model amid a slashed budget and high staff turnover. An EYEN staff member described programme implementation as a generally overwhelming process, while the CEO described taking a systemic view of the problem and initially finding it overwhelming. I observed Heather to be

overwhelmed by the messiness of EYEN's ToC development. Several factors increase the contextual complexity faced by each NPO, which may lead to team members feeling overwhelmed. These are discussed in the section below.

Sources of Contextual Complexity. EduSpark and EYEN both serve impoverished urban communities that have an urgent need for ECD services. A 2021 census of ELPs in South Africa found that there are, on average, six programmes per 1 000 children across the country (DBE, 2022), indicating a significant gap in universal access to ELPs. NPOs play a significant role in addressing this gap in access through ECD service provision and funding for the sector (Atmore, 2012, 2013; Horler et al., 2023). The organisations' sense of obligation to meet their communities' needs was evident in the SAMEL workshops, particularly amongst the EduSpark team. In the post-test follow-up interview, EduSpark's CEO described the ECD community as *"anxious and jittery"* amidst the uncertainty that their community was facing, as EduSpark had lost much of its ECD funding and personnel, curbing their ability to sustain service delivery in the community. Following the SAMEL workshops, the EduSpark team held an information session with ECD principals, to present the new programme model designed in the SAMEL workshops and *"to assure them that we, the programme, is stabilised"*, demonstrating the organisation's commitment to meeting the community's needs.

In the face of overwhelming need and resource constraints, the NPOs face tough decisions about what they can do with what they have to best serve their communities. Workshop discussions centred on defining programmatic boundaries often led to debates on what should be done in light of needs versus what can be done in light of personnel, resources, and the implementation environments. The EYEN MEL manager summarised this in the post-test follow-up interview:

“We can't do everything. We want to; we wish we could. We'd want to fix the whole world and wrap it in a bow. But, I mean, it's not realistic, and it's not a reality in the context that we are in.” (EYEN MEL manager, follow-up interview)

As a complete discussion of all sources of complexity affecting NPOs is beyond the scope of this thesis, I focus on the two most prevalent sources that continuously surfaced throughout the SAMEL tests.

Firstly, South Africa's ECD policy framework significantly affected both teams' conceptualisation of their programme designs. The previous chapter described several instances in which workshop discussions centred on policy-related issues, which often instigated debate and led to frustration amongst the teams.

The most contentious issue that continuously resurfaced throughout the workshops was that of ELP registration with the provincial government. As explained previously, if ELP facilities register as partial care facilities, they receive a government subsidy per child, per month. To be eligible for registration, facilities need to meet several criteria, satisfy provincial norms and standards for health and safety, and submit a portfolio of relevant documentation. While registration was not the core programmatic objective for either EYEN or EduSpark, both organisations grappled with its place in their programme design, given the organisations' aim to support the quality and sustainability of early learning opportunities in their communities.

For example, the successful registration of ELPs originally formed part of EYEN's ToC. In the ToA workshop, the team debated the relevance of registration as an outcome. I recorded these comments during the discussion:

“Mandla points out that the ECD registration outcome isn't applicable to all their ELPs because many of them can never be registered, despite the model being the same for all. He says: ‘For me, that's a flawed outcome’. Gail agrees, saying that their bigger

goal is quality, rather than registration, which Mandla agrees with. He adds: 'Registration is a great achievement because it unlocks funding ... but the main outcome is not a tick-box that a centre is registered. There are so many subcomponents. I don't want to overcomplicate things, but I think we should revise that outcome.'" (EYEN ToA observation notes)

The EduSpark team emphasised ELP registration more in their ToC and programme design, deciding to support ELPs that can register. Consequently, this raised questions about the programme's relevance and place in supporting ELP quality. In the ToA workshop, the CEO raised concerns about the motivation for ELPs to engage with EduSpark, wondering if they would merely participate to achieve compliance for registration. They noted that EduSpark was failing to provide ongoing support to ELPs, and that the ToA workshop highlighted critical gaps in the implementation of their strategy to ensure ELPs deliver quality services.

In the previous chapter, under Theme 1, I described the power of boundary critique as a discursive tool that helped the teams deliberate on their target populations: ELP participants who can register versus those who will never be able to satisfy the eligibility criteria. These conversations highlighted profound policy issues that ECD organisations contend with in the early learning landscape. Should programme resources be focused on ELPs that can register, or informal ELPs that cannot but serve an otherwise untouched portion of the child population in poor areas? What happens to children who cannot afford ECD centre fees and have no other means to access early learning opportunities? As described earlier, these were contentious areas of debate in both tests.

A second major source of complexity for EYEN and EduSpark was the role played by donors in their programme design and MEL requirements. This section focuses on

programme design, while MEL requirements (Theme 3) are discussed in the subsequent section.

The issues of donors surfaced in all engagements with the NPO teams, except the Scoping workshops. As noted in the previous chapter, the EYEN team repeatedly grappled with pressure from donors to include child outcomes in their ToC, even though they did not work directly with children. Both teams also described implementing donor-led intervention models to fund their programmes via grants. Recall this extract from the EduSpark ToC workshop:

“We’ve been implementing a very haphazard ECD model that was very much wide stretch and actually donor-led.” (EduSpark programme manager, ToC focus group)

Both teams described the tendency to design interventions to align with donor interests in order to secure funding, leading to programme models that are predominantly donor-led instead of community-driven. As ToCs are typically incorporated into funding applications, the teams debated whether they should be aspirational and impressive in their ToCs, or more constrained, in line with their current budgets (and then possibly appear less impressive).

“Are we not dreaming too much?” (EduSpark ECD programme manager, focus group)

The sources of complexity discussed above had important implications for the effectiveness and feasibility of the SAMEL process. This is discussed in the following section.

Effects on Programme Theory Development and the SAMEL Agenda. As described in the previous chapter, explicating each organisation’s core interventions was not an easy task for me or Heather. Perhaps due to their government partnerships, both organisations seemed to struggle to conceptualise their programme in light of ECD policy and donor interests. This

was exacerbated by the fact that both teams described coming together in strategic engagements, such as the SAMEL process, as a rare occurrence. This is addressed in greater detail in the discussion on Theme 3. However, it is important to note here the effect on programme theory development: the SAMEL test appeared to be one of the first instances in recent history in which the management teams in both organisations revisited their programme designs. Consequently, discussions about programme design were long and sometimes contentious (especially around the policy issue). I had not accounted for this in the Toolkit's design, expecting that two workshops (ToC and ToA) would be sufficient. They were not.

In both the Evaluation and the Monitoring & Learning workshop, teams were still grappling with programme design issues, which significantly limited the relevance and feasibility of the monitoring strategy and evaluation plans that were being developed. The EduSpark Monitoring & Learning workshop revealed important gaps in the programme interventions that the team had not discussed previously, leading to a significant amount of time being dedicated to getting the team on the same page. Amongst the EYEN team, there appeared to be glaring gaps in individuals' knowledge and understanding of the programme (recall the "collective meltdown" described in the previous chapter), which led them to revisit the ToC outcomes. In the final SAMEL workshop, an EYEN team member spotlighted the need for further collective reflection:

"What I would really like to see, following this, is how we roll this out. Because, internally, even up until, like, this last question ... people still have very different ideas of what EYEN's strategic goals and visions are... So, for me, I would love to have a conversation afterwards. Because at this level, ... we actually need to refine and define what EYEN is meant to be doing." (EYEN team member, Evaluation focus group)

In summary, complex contextual forces (particularly ECD policy and donors) influenced and shaped workshop discussions, leading to tension between a systematic theoretical workshop agenda and the complex, dynamic context in which the learning would be implemented. The workshop data highlight the advantage of systems thinking concepts in surfacing and managing this complexity productively, but, on the other hand, this complexity significantly disrupted the SAMEL agenda, due to the time needed to discuss and explicate an agreed-upon programme model.

Theme 3: Organisational Disconnect

Yet another inescapable reality of non-profit work is how external and contextual complexities seep into NPOs' organisational structure and culture. In the SAMEL tests, it was evident that both EYEN and EduSpark were characterised by organisational features that their contexts had shaped, which notably constrained the organisations' readiness for MEL development. As the NPOs are mandated to provide critical services in resource-starved contexts. Having limited resources themselves, many organisational endeavours are seen as luxuries, such as engaging in reflection and strategic work, or hiring MEL personnel. This theme highlights an important benefit of the SAMEL process: an opportunity for team members to collaborate and engage in collective strategic work (as discussed in Theme 1). However, it also identifies several barriers that limited the effectiveness of the workshops, discussed below.

Barriers to MEL Readiness. The test data revealed a few features in both EYEN and EduSpark that indicated low levels of MEL readiness. In this context, I define MEL readiness as the extent to which an organisations has the resources and capacity to design, implement, and maintain a MEL system, coupled with the institutional support to do so.

The first barrier that I identified early in the test data is NPO teams' need for **more time** to engage in strategic work. Strategic engagement is considered a luxury, and is overshadowed by operational duties. EYEN's managing director noted this in a workshop focus group:

"We haven't had the luxury of time to think strategically to build up these models, because everything is operational, you know?" (EYEN managing director, ToC focus group)

Having limited time to engage in processes such as the SAMEL workshops was noted by both teams. For example, the EduSpark fundraising manager expressed a trade-off between sacrificing time to engage in strategic engagement versus the benefits of doing so:

"Most organisations short-circuit this process purely because of time constraints. It's not everybody that can set out six, seven sessions or the whole day to thrash this out ... Because, I mean, I'm thinking just how heavy it was on us to have a day out of the office, and again tomorrow. But it is necessary. That shortchange or that trade-off really shouldn't happen. It shouldn't happen." (EduSpark fundraising manager, Monitoring & Learning focus group)

As noted in the discussion on Theme 2, the consequence of limited time together meant that the SAMEL workshops were the first time in recent history that management teams could come together to discuss strategic issues. This was seen as a significant benefit by both teams. However, this meant that the workshop agenda had to be pivoted in many instances to accommodate much-needed conversations. As both the workshop facilitator and Toolkit designer, I found this frustrating, captured in this journal extract:

"I also get frustrated when they go on tangents, or Dion asks them random questions about how they used to do things ... This is a coming together of the team, and it feels

like an opportunity for them to talk through things they wouldn't normally, which is amazing, but — Hello! We have a very specific agenda.” (ToA journal)

One of the ways in which the NPOs appear to manage their operational load in such resource- and time-limited environments is by operating programmes in silos, particularly EYEN, a far larger organisation than EduSpark. Siloed programme management is seemingly intended to increase organisational efficiency, as the EYEN CEO explained:

“We've got five different programmes in this organisation. And, as an organisation, it's hard, because it's pulling your resources and competing with different things. But we need to almost look at it as if we were five different organisations, and actually ringfence and support, and actually have those teams. Because me, as a CEO, I can't run five organisations, but we can have five CEOs run five organisations.” (EYEN CEO, Evaluation focus group)

A consequence of these silos, combined with limited time for strategic engagement, is the tendency of teams to be strategically misaligned. Quoted in the discussion of Theme 2 was an EYEN team member who, in the final SAMEL workshop, noted that the EYEN team *“still have very different ideas of what EYEN's strategic goals and visions are”*.

A second consequence, mainly due to time constraints, is a **lack of a learning culture**, which was described by both organisations across several focus groups. The EYEN managing director expressed a lack of time to create institutional learning structures due to operational demands:

“But, unfortunately, you've got operational responsibilities, which just trump all of this luxury of having to put together these processes and system and time.” (EYEN managing director, Monitoring & Learning focus group)

Both EYEN's and EduSpark's CEOs were particularly enthusiastic about the need for a learning culture in their organisations, but, in the follow-up interviews, they expressed that learning strategies had been neglected following the SAMEL tests, due to other operational priorities.

As with strategic engagement, a learning culture was seen as a luxury, without a strong motivating force to develop such a culture. This was particularly evident in the ways in which both NPOs conceptualised MEL — primarily as an externally motivated accountability mechanism, rather than an internally motivated mechanism for learning. For example, the EYEN managing director framed the importance of a monitoring framework in light of donor accountability:

"I think just having that framework to be able to present from a funder- and evidence perspective is important". (EYEN managing director, Monitoring & Learning focus group)

Similarly, EduSpark's fundraising manager related the learning strategy to donor requirements, suggesting that much of the organisational motivation for MEL stems from donor requirements, as opposed to an internally-driven motivation to learn:

"Timothy mentioned that the following question was beginning to pop up from funders: 'Are you a learning organisation?'" (Observer notes, Monitoring & Learning workshop)

The role of donors' reporting requirements was an ongoing issue during the SAMEL workshops. Both teams repeatedly referenced these requirements when developing their MEL frameworks, attaching greater weight to the requirements than the kind of data that would be important for learning and reflection. Both teams described limited agency over what data needed to be collected.

“We understand the nature of the reports and having to deal with thirty different funders on a weekly basis. That's the info they want to see. And, unfortunately, that's what we need to provide.” (EYEN managing director, ToC workshop)

“Certain donors may come with certain indicators ... That's often a mistake that is made. 'So okay, they want this, take this out and put that in.' It should never happen that way.” (EduSpark fundraising manager, Monitoring & Learning focus group)

As NPOs typically receive funding from several donors, which donors have different reporting requirements, MEL systems risk becoming *“fragmented”*. In other words, multiple donor requirements lead to disjointed MEL and reporting systems, as noted by the EduSpark fundraising manager:

“I think we've probably suffered ... with M&E sections being fragmented, because that's what the donor wanted. But now you're sitting at something that ... it's not coming together properly.” (EduSpark fundraising manager, Evaluation focus group)

In the follow-up interview, the EYEN MEL manager reported using the work done in the SAMEL workshops to create the logframe required by the organisation's government donor. Both CEOs revealed that the Scoping and Focusing work was beneficial for funding applications, and both teams had used this work to attract donors, demonstrating the motivational role that funding plays in an organisation's drive for MEL.

Finally, a significant barrier to MEL readiness that was observed in the SAMEL tests is a general **lack of institutional capacity for MEL**, particularly in terms of personnel, training, and organisational buy-in. While EYEN had a dedicated MEL manager, EduSpark did not. The lack of personnel capacity to hold the MEL function was noted across the EduSpark workshop focus groups as the team attempted to design a comprehensive MEL framework. They struggled to find a balance between a framework that seemed out of reach and a significantly

watered-down version that was in line with their current capacity. This tension was expressed in the Monitoring & Learning focus group:

“So, I think maybe that's the challenge ... Do we have the resources and the capacity? And yet, we can't shortchange what needs to be recorded simply because we don't have the people. What we need to do is find a better way to do it - to get us there.”

(EduSpark fundraising manager, Monitoring & Learning focus group)

Despite having a MEL manager, the EYEN team felt similarly ill-equipped to undertake comprehensive MEL activities with a one-person MEL department, who was new to the organisation. At times, the MEL manager seemed overwhelmed by the workload. This was evident in a comment made in the Monitoring & Learning focus group:

“Because, in my head, you tell me we need to monitor 698 [ELPs]. I'm gonna be chopped in pieces, and you can post me to all the centres to go collect information.”

(EYEN MEL manager, Monitoring & Learning focus group)

Ensuring organisational buy-in and training in MEL were noted as vital mechanisms to support low levels of capacity in both organisations under study. In the post-test follow-up interview, the EYEN CEO noted an ongoing need to instil the MEL function across the entire organisation. All team members would know their role in implementing the MEL framework, so as not to put the burden on a single team member. In the EduSpark follow-up interview, the CEO also noted the need for a greater understanding of the MEL framework across the organisation. The CEO was also concerned about a lack of understanding of what MEL involved amongst employees who had not participated in the test.

In summary, the EYEN and EduSpark test data revealed several organisational barriers to MEL readiness, including a historical lack of time for strategic engagement (due to operational priorities and siloed programme management), that disrupted the SAMEL

workshops. Furthermore, limited institutional capacity for MEL (particularly in terms of personnel and training) and a historically fragmented and donor-led approach to MEL further eroded the teams' institutionalisation of MEL as an organisational learning aid to enhance programme effectiveness. Instead, MEL was viewed primarily as an accountability tool, with little priority given to embedding MEL throughout the organisation. This hampered the SAMEL process, as the teams did not appear ready for the design of a comprehensive MEL system, nor did they appear to have the required institutional capacity to implement such a system. The SAMEL workshops resulted in MEL frameworks that, according to the follow-up interview data, appeared to be a long way off from being implemented.

Relationships between Themes

As described in the introduction to this chapter and illustrated in Figure 53, the themes presented here are not linearly related. The + and – signs again indicate reinforcing and balancing feedback relationships between the themes, while the numbers pertain to the four relationships illustrated in Figure 53. Drawing on the findings of the thematic analysis, the following claims are hypothesised.

- 1  If organisations have low levels of MEL readiness, characterised by limited strategic engagement, fragmented MEL systems, and strategic misalignment amongst team members, a process such as the SAMEL workshops could have a powerful positive influence on increasing MEL readiness through enhanced strategic clarity. The Scoping and Focusing workshops could enhance teams' ability to deal with their complex contexts, consider their programme scope and boundaries, and collectively define focus areas.

⊖ However, low levels of MEL readiness can negatively affect the effectiveness of the Scoping and Focusing workshops for such teams, due to the amount of time these workshops require. Developing programme theory can be a lengthy and messy process if teams are coming together for the first time to address design issues.

2 ⊕ The strategic clarity gained in the Scoping and Focusing workshops could enhance teams' ability to strategically engage with the complex contexts in which they work. The systems thinking concepts introduced in these workshops could help teams focus and bound their programmes effectively.

⊖ However, the contextual complexity of the programmes also complicate the development of programme theory. Sources of this complexity, such as ECD policy and donor pressures, complicate the design decisions the teams need to make to conceptualise their programme theories.

3 ⊕ The contextual complexity surrounding organisations influences the organisational structure and resources needed to develop a feasible MEL framework. A reinforcing relationship can be hypothesised here, whereby the more complex and demanding the context and community needs are, the greater the operational demands on the organisation, leading to siloed operation and limited time allocated to strategic engagement. As a result, organisations may use MEL primarily as a donor accountability tool, leading to a lack of a learning culture.

4 ⊖ In turn, this organisational disconnect erodes organisational readiness for MEL and undermines the feasibility of a five-workshop MEL development process.

In conclusion, the feasibility of implementing a systems thinking approach to MEL

development in NPO contexts, as explored through the SAMEL Toolkit, presents both

promising opportunities and significant challenges. The reflexive thematic analysis findings indicate that systems thinking substantially enhances programme theory development, which, in turn, aids organisations in programme design and strategic planning, setting the stage for enhanced MEL readiness. However, the findings also demonstrated notable contextual and organisational barriers inherent in non-profit settings, particularly related to organisational capacity for MEL and accountability structures required by funders. These factors constrained the SAMEL Toolkit testing process, particularly in terms of having enough time to finalise the programme theory and translate the theory into a feasible MEL framework that is aligned with available capacity. Thus, the tests revealed the real-life constraints that can shape the effectiveness and success of designing and implementing a systems-based MEL framework. The implications of these findings are discussed in more detail in the following chapter, in which I present a discussion of all the test findings and make recommendations for further research.

Chapter Ten: Discussion

This exploratory research aimed to develop and test a systems thinking Toolkit for MEL development in ECD NPOs. The research used a sequential two-phase design, in which the toolkit was first designed and reviewed by experts and then field-tested with two ECD NPOs. I employed a qualitative orientation to capture my nuanced experiences, as well as those of Heather and the participating NPO teams. This chapter presents a discussion of the major findings pertaining to the test phase of the research. Drawing on the preceding two chapters, this chapter consolidates the test description and thematic analysis findings, highlighting three major findings on the Toolkit's feasibility in real-world settings. Implications of the applicability of systems thinking to MEL, as well as directions for future research, are also discussed.

The Approachability of the Toolkit for Evaluators

The SAMEL Toolkit was designed with the following quote about systems thinking at the back of my mind:

Any field requires fresh minds and next generation adoption, and this is very relevant for systems thinking at the present time, given the highly complex global-to-local problems that we are experiencing, where systems approaches are sorely needed. The Cambrian explosion of frameworks made the field fragmented and unapproachable. As the field evolves, it needs to be welcoming and embrace newcomers (unlike the 'weed out' often seen in the cultures of physics, chemistry, biology, and medicine). In order to do so, systems thinking must be explained succinctly, and people must see its value readily. (Cabrera et al., 2023, p. 16).

I spent a significant amount of time and effort trying to make systems thinking approachable in the SAMEL Toolkit, drawing on my own experiences as a newcomer to the

field, as well as aesthetic principles of instructional design (Parrish, 2009). The Toolkit uses simple language, provides many examples, and employs graphics, colour-coding, and worksheets to make the concepts approachable, user-friendly, and digestible. Additionally, I opted to use the systems thinking concepts approach in the design of the content, focusing on systems, relationships, boundaries, and perspectives (discussed further below).

The test findings revealed some success in ensuring approachability of the Toolkit. Heather and the expert reviewers regarded the format and presentation of the Toolkit as approachable and user-friendly. Heather also reportedly gained an appreciation for the conceptual benefits that systems thinking can offer MEL. However, she also demonstrated limited internalisation of all of the Toolkit's systems thinking tools and underlying methodologies by prioritising workshop preparation and not an in-depth understanding the content. Instead of reading the SAMEL chapters cover to cover and then starting workshop preparation, Heather reported copying the content into her workshop slide deck. As discussed in Chapter Eight, this appeared to be a consequence of time limits; due to Heather's consulting workload, she needed to prioritise, in her words, "*getting the job done*". Time played a significant role in her engagement with the process — a critical factor in the real-world feasibility of the SAMEL Toolkit. In her final interview, Heather said that time was her biggest barrier to engaging in professional development or upskilling activities.

Heather did, however, facilitate some SAMEL concepts and activities particularly well: the SEM (Bronfenbrenner, 1979), boundary critique, actor mapping (Gopalakrishnan & Clarke, 2015), the partner analysis for the ToA, and the service utilisation diagram (Rossi et al., 2019), some of which she described as her favourites and ones she could see incorporating into her professional practice. These were also my favourite tools to facilitate, as they were easily comprehended, and appeared to resonate strongly with team members. On the other hand,

Heather did not view CATWOE (Checkland, 1981), CLDs, system archetypes (Senge, 1990), and CSH (Ulrich, 1983) favourably, as she found them confusing and not particularly helpful in MEL development (although she noted that CATWOE was valued by the EYEN team).

It is possible that Heather's least favourite tools have theoretical underpinnings that require more time to learn and internalise. For example, the service utilisation diagram (Rossi et al., 2019) and actor mapping (Gopalakrishnan & Clarke, 2015) are relatively straightforward exercises that do not require much additional reading to grasp, whereas CATWOE (Checkland, 1981) is underpinned by an extensive methodology (SSM; Checkland, 1981), which may not be adequately introduced by the SAMEL Toolkit. This may also be the case for CLDs and system archetypes (Senge, 1990) (underpinned by system dynamics), as well as CSH (Ulrich, 1983).

Thus, there was variability in the extent to which Heather was able to internalise and strongly facilitate some concepts, as well as the extent to which Heather 'bought into' the systems thinking concepts, possibly determined by the confluence of the tools' degree of complication and Heather's available time to engage with them. For example, as described in Chapter Eight, Heather said she did not understand CLDs, nor did she have the time to study them. This would suggest that some systems thinking methodologies that underpin SAMEL tools, like system dynamics and SSM (Checkland, 1981), require a clearer introduction to ensure that their core tenets are not lost on workshop participants, as they may not have time to engage in additional reading or learning. While I struggled to facilitate CLDs and system archetypes (Senge, 1990) in the workshops, I resonate strongly with the concepts, and intend to keep practising and incorporating them in my future consulting work. This is likely because I, unlike Heather, had time to read up on and acquaint myself with them.

Both Heather's and my experience lend credence to the idea of the systems concepts as a "useful entry point" (Walton et al., 2021, p. 167) into the use of systems thinking in the evaluation field. We were both able to obtain an understanding of what systems thinking is (taking the view that the core systems concepts represent what a systems thinking lens is), and see its relevance to and utility in MEL. As described earlier in this thesis, the systems concepts were very helpful in my initial introduction to systems thinking and in the application of the Toolkit.

However, the test data revealed that several of the tools used in the workshops were watered down, or some of the meaning espoused by their underlying methodologies was lost. For example, Heather did not adhere to many of the core Gestalt principles of life spaces (Lewin, 1936), such as attending to the relationships between factors, the extent of each factor's influence, or the role of secondary factors. Likewise, I did not adequately attend to reinforcing and balancing relationships in EduSpark's CLD session, opting for a simpler discussion.

The misuse or misinterpretation of systems concepts or methodologies is precisely what critics like Jackson (2023; see also Miller, 2016) warn is the problem with a systems concepts approach to evaluation. While the SAMEL Toolkit attempts to sidestep this issue by providing explicit tools and instructions for applying the systems concepts (thus not leaving the interpretation of the concepts up to the evaluator), the findings of this research indicate that these steps were not sufficient to foster an in-depth application of all tools and their underlying methodologies. It is unlikely that this is because the Toolkit was not simple or approachable enough, based on Heather's and the reviewer's sentiments, but rather because too much has to be done in limited time. Most of the tools used in the workshops provoked rich, in-depth discussion, but both tests were regularly short of time because each

workshop's agenda was packed tightly with several systems thinking tools and activities. Further, the Toolkit requires evaluators to learn about a substantial number of tools and methodologies, for which they might not have the time.

Thus, further work is necessary to refine and test how a simpler, pared-down version of the SAMEL Toolkit would fare in similar settings. Based on the outcomes of the test study, it would be worth exploring replacing the Evaluation workshop with an additional workshop to finalise the programme theory.

Having explored the approachability and initial application of the SAMEL Toolkit by evaluators, I now turn to its strategic implications in MEL, focusing initially on the major benefit that was observed.

Systems Thinking's Major Contribution to MEL: Programme Theory

In the previous chapter, the theme *Strategic clarity (Programme theory)* described the major advantage of the SAMEL process for the NPO teams. The test data revealed the NPO teams' ongoing appreciation of a structured opportunity for strategic engagement, particularly with regard to developing a programme theory informed by systems thinking. The Scoping and Focusing workshops were the best-received of all of the workshops, as the teams found immense strategic value in the theory development process, particularly in the way it provided them with a collective sense of clarity and a direction towards structured programme design that was aligned to their environmental context.

Programme theory was used as the basis for MEL development in the SAMEL Toolkit because I agree with theory-based evaluation proponents (e.g., Chen, 2015; Donaldson, 2007; Funnel & Rogers, 2011; Rossi et al., 2019) on the usefulness of programme theory in explicating assumed causal relationships between programmes and their goals. The aim was to help teams test the assumptions underlying the logic of their programmes. In my

professional consulting practice, programme theory is a consistently powerful reflection- and planning tool for clients. Programme theories, and particularly ToCs, have been found to be helpful tools for prioritisation, deliberation on context, sharpening focus, planning, and monitoring and evaluating the outcomes to which the programme is contributing (Douthwaite et al., 2020; Mackenzie & Blamey, 2005; Valters, 2014; Vogel, 2012).

The NPO teams were highly responsive to the SEM (Bronfenbrenner, 1979), CATWOE (Checkland, 1981), actor mapping (Gopalakrishnan & Clarke, 2015), boundary critique, and life spaces (Lewin, 1936) systems thinking tools that were employed in the Scoping and Focusing workshops, which were also used to inform the development of a ToC. According to the test data, these tools had significant advantages, which are summarised in Table 9.

Table 9

Major Advantages of Scoping and Focusing Systems Thinking Tools

Concept	Major advantage
SEM	Understanding where the programme fits into the larger ECD ecosystem
CATWOE	Understanding the role of other actors in addressing the problem
Actor mapping	Identifying opportunities for partnership
Boundary critique	Reflecting critically on values of inclusivity and the implications of exclusivity in determining the programme population
Life spaces	Understanding the environmental context and its role in the programme's design and implementation

The systems thinking tools listed above helped the NPO teams to develop a shared understanding of their respective programmes, reflect on their contexts in a meaningful way,

focus on desired impacts and outcomes, make key intervention boundary decisions, identify future opportunities and stakeholder relationships, and identify implementation barriers.

Similar findings have emanated from other studies in which systems thinking was used to facilitate programme theory in complex programmes. For example, network mapping, a similar approach to actor mapping (Gopalakrishnan & Clarke, 2015), has been found to enhance theory development by fostering an appreciation for power dynamics in programme contexts and, as in the SAMEL tests, the need to collaborate to create change (Douthwaite et al., 2020). Incorporating multiple diverse stakeholder perspectives (as EduSpark did by including ECD principals) has been found to help NPOs conceptualise their role in partnerships that create change and a shared goal (Apgar et al., 2017). Other authors argue for the utility of identifying enabling contextual conditions (as with life spaces) to foster effective programme adaptation (Rice et al., 2020), as well as the conscious deliberation on programme boundaries (as in boundary critique) to surface values of power, inclusion, and exclusion (Arkesteijn et al., 2015). Finally, realist evaluation has been successfully used to develop programme theory for large, complex programmes (Shearn et al., 2017).

The test data showed that including non-linear feedback in programme theory was challenging. Chapman et al. (2023) looked at different forms of complex mapping in ToC development for multidisciplinary research projects, and found that systems-based mapping approaches (incorporating non-linear feedback) were rarely used, compared to more straightforward or linear approaches, which serve “as a ‘big win’ public relations, communication and accountability tool” (p. 319).

Had Heather and I had more experiential opportunities to develop our understanding of feedback loops and CLDs, and if there had been more time in the Scoping and Focusing workshops to engage with this material more meaningfully, these systems dynamics concepts

(feedback and non-linear relationships) may have generated greater meaning than they did in the tests. The Toolkit clearly conveying these concepts thus requires attention.

CLDs are referred to as 'horrendograms' by some, due to their propensity for paralysing stakeholders by "the sheer scale and messiness of complexity", rendering action seemingly impossible (French et al., 2023, p. 27). Feeling paralysed or overwhelmed by complexity and large-scale system representations was recorded in the EYEN test, suggesting that there is an important balance to strike in terms of accurately mapping intended social change versus the purpose and utility of the programme theory, as well as the ability of the theory developers (NPO teams in this case) to engage meaningfully with the theory during and after development (Chapman et al., 2023; Davies, 2004). If programme theories are to be owned and regularly used by NPO teams, they cannot be confusing or seem difficult to apply. The test data suggest that using CLDs or systems diagrams may have rendered the programme theories confusing and unhelpful to the teams, preventing ongoing engagement with the theories. This affirms the use of the linear diagram approach in the SAMEL Toolkit, alongside minimal use of feedback loops (to indicate the most important non-linear relationships).

The advantages discussed above and summarised in Table 9 call into question the overemphasis on wicked problems and complex contexts that currently dominate the systems thinking conversation in programme evaluation. In Chapter Three, I described the tendency of evaluators to view a situation as either appropriate for systems thinking or not, primarily relying on heuristics such as the Stacey matrix (e.g., Funnel & Rogers, 2011; Patton, 2011). This results in systems thinking being used only for programmes that are seemingly complex in design, or in aiming to achieve complex systems/policy change (e.g., Hargreaves, 2010; Latham, 2014; Parsons, 2007; Preskill et al., 2014; Urban et al., 2021; Wasserman, 2010). In

contrast, the SAMEL tests demonstrate that systems thinking can be successfully utilised by NPOs working with programmes that are not required to be completely adaptive or responsive to unpredictable and emergent change processes.

As argued in Chapter Three, child development is an inherently complex change process. However, it is primarily a sector that is well understood using empirical evidence. The EYEN and EduSpark teams were highly knowledgeable on the necessary inputs to create meaningful improvement in early childhood education, yet they still benefited from the systems thinking tools introduced in the Scoping and Focusing workshops (see Table 9). A systems-enhanced programme theory thus has immense strategic value for programme planning. However, there were also notable challenges to this approach to programme theory development, discussed below.

The Shortcomings of Systems-based Programme Theories

Most of the resources and methodologies I drew on to design the SAMEL Toolkit were rooted in systems thinking to understand situations. This is likely why most efforts that blend systems thinking and evaluation focus on the conduct of external evaluations (i.e., understanding programme outcomes or impact – e.g., Trochim et al., 2016) and ToC development (i.e., understanding programme logic – e.g., Chapman et al., 2023; Douthwaite et al., 2020), rather than MEL development, with the exception of Outcome Mapping (Earl et al., 2001) and Reflexive Monitoring in Action (Van Mierlo et al., 2010).

As a consequence, the SAMEL Toolkit is weighted towards programme theory. The Scoping and Focusing chapters and workshops feature the most systems thinking tools and concepts, while the MEL sections have less. The advantage of this, as evidenced by the outcomes of the tests, is an enhanced strategic approach to programme theory development.

However, there are also a number of drawbacks to this expanded approach to programme theory.

Firstly, it requires a deep engagement with complex concepts, which is a time-consuming discursive process that requires more time and organisational engagement than were afforded in the SAMEL tests. Programme theory questions and debates often consumed discussions in the subsequent workshops (Evaluation and Monitoring & Learning). The Monitoring & Learning workshops, in particular, were rushed in both tests, as programme theory discussions were still taking place. Secondly, as the programme theories were so comprehensive, the identification of outputs and indicators (traditionally an already time-consuming and laborious process) was complex and lengthy in the Monitoring & Learning workshop. As such, the SAMEL Toolkit's approach to theory-based monitoring may be unrealistic for two reasons.

It is Time-consuming. More time appears to be needed to finalise the programme theory and get organisational buy-in (and board approval, if necessary). The SAMEL test outcomes suggest that the Evaluation and Monitoring & Learning workshops took place too soon after the Focusing workshops, in which the ToC and ToA were drafted. The teams would have benefited from more time between workshops, as well as additional time to finalise the programme theory (as mentioned earlier). However, while more time was needed, it was not necessarily available.

The SAMEL test process was acknowledged as a time-intensive undertaking for both organisations, necessitating a commitment of five full working days by the organisations' senior management teams. Both teams considered this a necessary and valuable investment, but, at times, it presented challenges in ensuring the continuation of organisational operations. As described in Chapter Eight, several staff members could not attend the

workshops due to time commitments. Operational demands often influenced the planning and implementation of workshops, as many team members had to attend to urgent matters during workshop time. This was to be expected, as participatory MEL activities in development or non-profit settings compete with operational demands (Adams & Garbutt, 2008; Chaudhary et al., 2020). When asked about this issue, both teams said they liked the way the SAMEL test was conducted in terms of time, with the EduSpark team, in particular, noting (on several occasions) that they wanted to complete the MEL development process as soon as possible. Thus, while more time allocated to SAMEL workshops might be ideal, it is unlikely to be afforded in reality.

It Requires More MEL Capacity. The SAMEL approach to programme theory necessitates a comprehensive monitoring system that may be incongruent with typical NPO resources. For example, monitoring changes to the programme's context is a key focus area in the SAMEL monitoring plan, requiring additional capacity for data collection and analysis. The EduSpark monitoring plan, for example, contained 52 progress indicators for one of the three programmes that were focused on in the test. The indicators represent change in short- and long-term outcomes, the programme's context, participant characteristics, intervention delivery, and programme factors that support implementation. During the test, I was resistant to developing this many indicators, but the EduSpark team wanted to be as comprehensive as possible, saying they would pare it down at a later stage. Increased data volume has been noted by evaluation experts as a significant pragmatic barrier to using systems thinking in evaluation (Walton, 2016), indicating a need for stricter guidance around indicator development in the Toolkit.

Follow-up interviews in the months after the SAMEL test revealed that the NPOs needed to make more progress towards actioning the MEL frameworks and developing

learning mechanisms. A lack of time and capacity were cited as the barriers to moving their new MEL strategies forward, suggesting that both teams were not completely ready to action the frameworks that were developed in the workshops. Subsequent to the conclusion of the SAMEL tests, both Heather and I were asked by the EduSpark and EYEN teams to continue consulting with them on their other programmes, on a paid basis. This suggests that the organisations had the financial resources for MEL, but lacked the expertise human capacity to drive MEL forward. This indicates that, for the Toolkit to be effectively applied and utilised by the organisations, MEL readiness needs to be attended to in some way.

During the workshops, a number of organisational barriers to MEL readiness were consistently discussed and observed. Most notably, there was a need for more human capacity to own and drive the MEL role in the organisations. EYEN, a large organisation with several programmes, had only one (newly recruited) MEL manager responsible for all MEL activities. EduSpark had no formal MEL personnel, and relied on staff to perform additional data collection and reporting tasks. Organisational buy-in was also noted as needing more attention in both organisations, particularly in terms of team members understanding the value of MEL, its purpose in the organisation, and what constitutes quality data. These issues are widely noted in the literature on evaluation capacity-building (Buetti et al., 2023; Chaudhary et al., 2020; Wade & Kallemeyn, 2020). This raises the question: does a systems thinking approach to MEL require more capacity than traditional approaches?

This study cannot reliably answer this question as capacity was low in both test NPOs. The data is therefore unavailable to determine whether high levels of capacity would increase the feasibility of an approach such as SAMEL. However, in my experience, EYEN and EduSpark's MEL capacity is characteristic of many South African NPOs. It is, in my view, therefore more important to ask: how can NPOs be supported to develop baseline MEL

capacity that can support a systems thinking perspective? I explore this in the following paragraphs.

Evaluation capacity-building (ECB) refers to the development of an organisational system in which staff are capacitated to sustain quality programme evaluation and the utilisation of evaluation findings in an ongoing manner (Stockdill et al., 2002). The ECB literature commonly refers to a lack of human and financial resources and a lack of an internal drive for evaluation as common barriers to building organisational evaluation capacity (Buetti et al., 2023; Chaudhary et al., 2020; Labin et al., 2012; Preskill & Boyle, 2009). While not an explicit ECB initiative, the SAMEL process could be a foundational ECB step, as it provides the strategic framework upon which internal MEL processes and protocols could be built. Furthermore, the test data revealed an increased appreciation for MEL and organisational learning amongst both teams, suggesting that the SAMEL workshops did transfer some aspects of ECB.

In sustaining evaluation capacity, Wade and Kallemeyn (2020) note the importance of resources, as well as the role of organisational leadership in identifying and meeting evaluation needs, particularly 'champions' that drive the process. These are leaders who actively understand and support the purpose of evaluation, set organisational expectations, encourage team members, use evaluation data, incentivise evaluation, and allocate resources to it (Taylor-Powell & Boyd, 2008). Evaluation capacity is also thought to be sustained by the development of evaluative thinking in organisations, being afforded the time to see the benefits of evaluation for the organisation, and the alignment of evaluation to the organisation's mission and values (Wade & Kallemeyn, 2020).

While human and financial capacity cannot be increased by an ECB initiative, evaluative thinking and leadership can. Field et al. (2018) suggest successful strategies for

embedding MEL in particularly resource-constrained settings, such as integrating MEL into all aspects of the programme's implementation and linking data collection with existing programme activities, thus ensuring that MEL is not perceived as a standalone activity or role, but the responsibility of the entire team. As noted in Chapter Eight, the CEOs of both EduSpark and EYEN expressed the need for MEL to be taken up by all team members in the organisation. This may necessitate including MEL activities in team members' job specifications and actively involving team members in data collection for evaluation, such as interviewing programme participants (Field et al., 2018). This may allow team members to (a) gain rich insights into programmatic issues and (b) appreciate the insights gained and the associated benefits of MEL tasks.

In conclusion, the SAMEL tests outcomes indicate that systems thinking has immense strategic value for the development of programme theory, as found in other studies (e.g., Apgar et al., 2017; Douthwaite & Hoffecker, 2017). However, using this theory as the basis for MEL has practical implications. It necessitates more time to develop and finalise than was available in the current non-profit settings, and it is associated with greater data needs. This is aligned with the views of experts in using systems thinking in evaluation, who highlighted the time- and financial implications of the need for large data volumes and participatory practices, such as the SAMEL workshopping approach (Walton, 2016). Drawing on the ECB literature, it is possible that MEL capacity constraints could be ameliorated through the development of evaluative thinking and leadership, and strategically integrating MEL with programme implementation, including the creation of more formal structures for MEL conduct (such as incorporating MEL activities into job specifications) (Field et al., 2018; Wade & Kallemeyn, 2020). This would necessitate a significant pre-process to the SAMEL workshops, aimed at developing evaluative thinking and organisation-wide ECB. Doing so may

also help to increase the time afforded to the theory development process. No ECB activities formed part of the current iteration of the SAMEL Toolkit, but doing so is worth exploring in future iterations, particularly the effect that ECB might have on organisations' ability to action a MEL framework based on systems thinking.

As noted above, the test data indicated a perceived need amongst the EYEN and EduSpark CEOs for greater MEL buy-in and ownership in their organisations. One particular reason why this may have been limited in both teams is the organisations' RBM orientation to MEL. This is discussed below as I present the third major finding of this research.

The Incongruence of Systems Thinking and Accountability-led RBM

The outcomes described in the preceding chapters indicated that the MEL development process was characterised by a focus on donor accountability amongst both teams. As described in the discussion of the *Organisational disconnect* theme, both organisations had donor mandates for reporting, and demonstrated a historical approach, whereby MEL is viewed as an accountability tool, rather than a learning tool. This may be one of the reasons for the lack of organisational integration of the MEL function in EYEN and EduSpark; MEL may be seen as an administrative task for funders, as opposed to an integral learning function to support programme effectiveness (Chaudhary et al., 2020).

MEL as an accountability tool reflects a results-based programme management strategy, which was initiated in the 1990s by the United Nations to improve effectiveness and accountability (UNDG, 2011). As described in Chapter Two, results-based management RBM was instituted in South Africa's non-profit sector post-1994, as international donors insisted on greater accountability amongst NPOs in the post-apartheid era, using tools such as the Logical Framework (Mouton, 2010). Accountability to funders remains an integral issue for MEL, as funders play a significant role in South Africa's non-profit landscape. The majority of

ECD NPOs, such as EYEN and EduSpark, rely on private donor funding, compared to government funding and self-generated income, according to a South African survey (Horler et al., 2023).

Accountability pressures can motivate organisations to seek out evaluation and MEL support, but accountability might not be adequate to maintain internal motivation and interest in MEL implementation and utilisation beyond short-term reporting (Buetti et al., 2023). An example of this from the test data is that neither organisation had progressed substantially in actioning the MEL frameworks developed in the tests, as the teams had to complete funder reporting and grant proposals. Both teams had drawn on the Scoping and Focusing work to assist them in obtaining funding, but neither had not put any other monitoring, learning, or evaluation elements into place yet. The test data thus revealed two major ways that funder accountability and RBM affected the SAMEL process, discussed below.

Intervention Design: Actioning the ToC

ToCs are heralded as particularly useful tools for explicating organisations' assumptions about how they intend to effect change by creating a logical argument linking interventions to intended outcomes and impact (Funnel & Rogers, 2011). As such, ToCs are a particularly powerful tool for programme design, as teams are guided to critically reflect on the logic underlying their programmes. However, the context of theory development is likely to shape the extent to which critical reflection takes place (Valters, 2014). This was evident in the SAMEL tests as team members raised a number of critical issues with their intervention designs during (and beyond) the Focusing workshops, but funding relationships were noted as a significant constraint in making design decisions (as noted in the discussion of the *Contextual disconnect* theme in the previous chapter).

What follows is a discussion of the role of external pressures in intervention design by non-profit funding frameworks. The SAMEL tests revealed how grant funding and government mandates notably constrained the NPOs' ability to effectively use the insights gained during the Focusing workshops, in which the ToC and ToA were developed. Thus, while systems thinking may provide rich insights into effective programme design, the ability to transform this into actual changes to an NPO's programme design is highly questionable in light of funding relationships.

A notable example is the insight gained by both teams in the boundary critique exercise. After developing a ToC in the first Focusing workshop, the SAMEL Toolkit proposes a critiquing exercise in the following (ToA) workshop, in which teams are guided to critically reflect on the intervention boundaries they defined in the ToC, together with the implications of these boundaries. Inspired by CSH (Ulrich, 1983), the Toolkit prompts programme teams to discuss the consequences of their programme boundaries by not only attending to the individuals involved in the programme system (participants), but also those who are not involved but are nevertheless affected by the system (i.e., excluded community members) (Gates, 2018; Williams & Hummelbrunner, 2011). CSH encourages one to reflect on boundaries in two modes: the actual *is* mode, which describes the system as it currently operates, and the normative *ought* mode, which describes how it should ideally operate (Ulrich & Reynolds, 2010). The aim is to think about who *ought* to be included in the programme, based on the situation analysis conducted during SAMEL's Scoping phase, and then compare this to who *is* currently eligible for the programme. Thus, teams can think about the practical consequences of the programme's design, participant eligibility in particular, and which community members may be inadvertently marginalised by the team's boundary decisions.

As described in Chapters Eight and Nine, this led to powerful and rich discussions in both tests; however, none of the insights generated in this session were actioned in either NPO's set of interventions. Despite naming populations that *ought* to be included in the programmes via another set of interventions, the teams did not include these in their final ToCs. Instead, teams focused on their original set of interventions, using the ToC process to describe *what currently is* as opposed to *what ought to be*.

Why was this the case? Both EYEN and EduSpark are government partners, and have historically been mandated and funded to provide support to ECD centres in their geographic areas of operation. National ECD policy explicitly outlines the critical role played by NPOs in ECD service delivery and in supporting the capacity of early learning practitioners (RSA, 2015). The Western Cape government began partnering with NPOs to support ECD centre registration in order to free up social workers to focus on community work (South African Government, 2018). As such, objectives for both EYEN and EduSpark include supporting ECD centre quality, supporting centres to comply with regulations and achieve registration with the DBE, training and mentoring ECD teachers and principals, and providing educational resources and infrastructural support. Despite the insights gained during the boundary critique exercise, in terms of the need for support amongst non-centre-based (or informal) ELPs, and the need to support children who cannot afford ECD centre fees (and have no access to non-centre-based provision), neither NPO could deviate from its core centre support interventions. This caused frustration amongst some team members, as there seemed to be a misalignment between the need to implement policy and the need to meet community needs. While non-centre based programmes are argued to be a cost-effective way to scale early learning delivery (e.g., Desmond et al., 2019), this area receives far less funding

by both local and global ECD donors as ECD centres remain a central government concern (Horler & Biersteker, 2022).

As such, ECD policy can be perceived as what Dahler-Larsen (2018, p. 7) refers to as a 'Janus variable' — an ambiguous variable that plays two (or more) roles in programme theory. A Janus variable can be functionally ambiguous when it has a positive causal effect in one relationship but a negative effect in another. For example, ECD policy positively affects formal ECD centres through subsidisation, but negatively affects non-centre-based ELPs that cannot register by marginalising them by withholding financial support. Both EYEN and EduSpark opted to 'park' the issue of supporting informal ELPs or providing financial support for children to attend formal ECD centres (raised in the sessions on boundary critique). For both organisations, this was a matter of the government's mandate, funding, and capacity. There did not seem to be an option to deal with the ambiguity of this particular Janus variable. Dahler-Larsen (2018) acknowledges that the extent to which ambiguity can or should be highlighted in programme theory depends on situational factors, such as the political interests undergirding programme theories. The author suggests evaluators take on an informational advisory role, pointing out the ambiguity and letting stakeholders decide how they want to deal with it. This is the course of action that Heather and I took, enabled by boundary critique.

In their review of ToCs in international development, Valters (2014) found that accountability and funding can have a "corrosive effect" (p. 12) on the honesty and usefulness of critical reflection in ToC development, as organisations can feel pressured to present their work in ambitious ways in order to align with government and funders' interests. Similarly, Apgar et al. (2017) found that many programmes prioritise accountability to donors over learning. The ToC process can thus be particularly uncomfortable when teams are confronted

with the need to adapt interventions in ways that could put them at odds with government partners (Apgar et al., 2017). This was evident in both SAMEL tests, as EYEN and EduSpark prioritised interventions that were aligned with national ECD policy (RSA, 2015), local government priorities, and funding opportunities. This limits the utility of the Scoping workshop, in which teams develop a situation analysis to which their programmes should respond, as their ability to respond is pre-constrained.

In terms of funding, both NPO teams reported historically designing their interventions according to grants and government funding. It is in the financial interest of EduSpark and EYEN to focus on ECD centre support, due to the related funding opportunities. In a survey of South African and international ECD funders, ECD centres were found to be the best-funded ECD service, focused on by 61% of survey respondents (Horler & Biersteker, 2022). This significant funding helps to explain why 54% of ECD NPOs in South Africa provide compliance- and registration support to ECD centres (Horler et al., 2023). On the other hand, during boundary critique, EYEN team members raised the issue of extending access to children who cannot afford ECD centre fees. In the ECD funder survey mentioned above, promoting child attendance at ELPs was rated as a top priority by 66% of funders, but was funded by only 20% of funders. There is thus a presumed lack of opportunity for NPOs, such as EYEN, to access funding to subsidise children's school fees.

The tendency of NPOs to align their programme strategies with donor mandates is well documented in the local and global literature. This is therefore not an issue unique to the ECD sector. Several authors describe donors' conditional grants as critical determinants of NPO strategy; NPOs shift their operations to align with donor mandates in order to secure much-needed funding (Bornstein, 2003; Heiss & Kelley, 2017; Reith, 2010; Witesman & Heiss, 2017). An ethnographic study based in Ghana described how NPOs develop grant proposals

that fit “with foreign donors’ aims, which in reality may be different from the needs of the local beneficiaries” (Yeboah, 2022, p. 1652). This finding was echoed in a qualitative South African study involving 10 NPOs across several social sectors. The authors found that “keeping donors happy” (p. 1821) is a top priority for survival, with some organisations becoming agents of donor programmes instead of implementing their own (Maboya & McKay, 2019).

The external pressures that impact intervention design thus significantly constrain the potential effectiveness of systems thinking in critically appraising NPOs’ chosen models of intervention in ToC development. Due to existing mandates and funding opportunities, NPOs may be unwilling or unable to action the critical reflections raised in the SAMEL Focusing workshops. The effectiveness of the workshops is therefore challenged by the NPOs’ lack of autonomy in their intervention design. While these trends do not cancel out the benefits of a systems thinking approach like SAMEL, they do highlight the critical relationship between MEL processes and NPOs’ funding landscapes, which evaluators that advocate systems thinking need to be conscious of and able to navigate.

Becoming a Learning and Adaptive Organisation

The second consequence of an RBM and accountability emphasis in MEL is the extent to which NPOs can become learning organisations. As noted earlier, there was great interest amongst both teams in becoming learning organisations, and the learning segment of the Monitoring & Learning workshops was received with interest. However, the organisations’ RBM orientations meant that MEL was not embedded in the organisations, but seen as an administrative task related to funding.

A notable critique of RBM is that it oppresses adaptive learning, stemming from its relationship with donor accountability. To avoid risking their relationship with donors, measures used to indicate NPOs’ performance can be misused to distort organisational

achievements (Mayne, 2007). RBM measurement frameworks are particularly vulnerable to this problem, as they tend to primarily emphasise short-term, measurable results, rather than longer-term, complex change (Simister, 2017). For example, a survey of ECD funders in South Africa reported that nearly half of the funders (49%) used reporting primarily for compliance and due diligence, rather than learning and iteration, and that the majority of grants have a duration of three years or less (Horler, 2021).

Compliance reporting within short grant periods incentivises NPOs to rigidly target short-term change rather than more complex and emergent change trajectories, which are harder to control and measure but could deliver beneficial long-term change (Apgar et al., 2017; Douthwaite et al., 2020; French et al., 2023). The MEL development process in both tests was dictated mainly by existing donor requirements and, thus, a focus on short-term change at the level of ECD centres. This was left largely unchallenged by Heather and me for two reasons. (1) Due to the MEL capacity constraints within each organisation, there is perceived pressure to focus resources on the most critical aspects for organisational survival. In the context of RBM, this becomes the outcomes that funders expect to see. (2) Due to a lack of workshop time during the development of the monitoring frameworks, we again had to focus discussions on these critical outcomes, as opposed to dedicating time to building mechanisms to monitor other emergent processes, such as unexpected outcomes or changes to the programmes' context. In settings with severely constrained resources and time, as well as expectations of RBM, these components of systems thinking feel like luxuries.

Many argue that, in the worst scenario, funders who enforce rigid RBM frameworks encourage organisations to 'game the system' by skewing, distorting, or falsifying programme performance data (Bornstein, 2003, 2006; French et al., 2023; Whitty, 2013) or even adapting or manipulating programme delivery to provide particular evidence to funders (Yeboah,

2022). To repeat some overused phrases in this regard, ‘what is measured is what matters’, as organisations focus efforts on seemingly easy wins, leading to a situation in which they ‘hit the target but miss the point’ (Bevan & Hood, 2006). For example, the fact that ToCs are used for both communication and accountability appears to affect the degree to which organisations use the approach to critically reflect on their policy and practice (Valters, 2014). In other words, “accountability trumps learning” (Whitty, 2013, p. 12).

In conclusion, while the SAMEL Toolkit was designed to address many of the criticisms of RBM and accountability-based MEL, the tests did not appear to achieve this objective. The requirements by EYEN’s and EduSpark’s funders simply did not allow for a flexible MEL environment that is fully conducive to actioning the insights gained from systems thinking. Evaluation commissioners and evaluation circumstances have been noted to constrain the use of systems thinking by evaluators (Gates, 2017). On this, Gates (2017, p. 167) argued the following:

To move towards fully taking up some of these approaches, there may need to be changes to the way evaluations are commissioned and contracted. One major consideration involves the kinds of social problems and change processes commissioning agencies aim to address. There are several trends in the types of initiatives in which evaluations are being commissioned that are particularly suited to the use of systems approaches (e.g. wicked social issues, social networks).

The present research thus adds evidence to the argument that changes are needed in how evaluation and MEL are commissioned, particularly by funders. However, the findings show that it is not just programmes addressing wicked problems that could benefit, as Gates (2017) also proposes. I argue that all social programmes would benefit from the strategic holism that systems thinking instils.

How Feasible is a Systems Thinking Approach to MEL Development in NPO Contexts?

Conclusions and Directions for Future Research

The findings of this study challenge the feasibility of the current iteration of the SAMEL Toolkit in terms of its comprehensive approach to the development of evaluator competencies in systems thinking, as well as the feasibility of a systems thinking-informed MEL framework for NPO settings. The implications of these findings are relevant in four main areas: (1) evaluator competencies; (2) theory-based MEL development; (3) NPO MEL capacity; and (4) the NPO funding landscape.

Firstly, the SAMEL Toolkit was partially effective in making systems thinking approachable. The tests provided evidence for an enhanced approach to using the systems concepts in evaluation by matching the concepts to systems thinking methodologies and particular tools that provide clear guidance on how the concepts should be interpreted and applied in a MEL setting. I maintain that this approach is one way in which Jackson's (2023) 'golden mean' may be approached in systemic MEL design. However, this necessitates a revision of the content of the SAMEL Toolkit to present a simpler approach that evaluators have more time and space to internalise, and that affords more workshop time with teams to finalise their programme theories.

Secondly, systems thinking proved to be well-suited to programme theory development in the tests. Programme theory was notably enhanced by particular systems thinking tools that enabled the critical reflection of relationships with other actors, boundaries of the programme, and the programme's context (see Table 9). However, this appeared to be a double-edged sword, as the process fostered comprehensive programme theories that necessitated equally comprehensive monitoring frameworks. This process was severely constrained by the five-day workshop format, which limited the amount of time

available to fully finalise the programme theory, as well as the time available to adequately engage with a monitoring framework for all the elements. This indicates that more time is needed to engage with the programme theory before the subsequent workshops on MEL can be meaningfully engaged with. These elements might be better suited to take place later on in the MEL development process, once there is organisation-wide buy-in for the programme theory, and some level of ECB has been developed (discussed below).

Thirdly, the NPOs in this study were limited in their organisational capacity for MEL. The systems thinking approach to MEL proposed by SAMEL adds significant complexity to ongoing monitoring activities, requiring the ongoing measurement of implementation, outcomes, and context. The ECB literature suggests that this is not necessarily an intractable problem, but that the development of evaluative capacity in organisations, particularly amongst leadership, may help to develop a greater internal drive for MEL, and thus direct the time and resources necessary to MEL tasks (such as theory development and MEL design, data collection, and learning events) (Field et al., 2018; Wade & Kallemeyn, 2020). For example, the SAMEL workshops could have been preceded by a thorough diagnostic process for ECB (Buetti et al., 2023), followed by a tailored ECB initiative to prepare each team for the MEL development process (Taylor-Powell & Boyd, 2008). The SAMEL Toolkit can be conceptualised as a supportive structure for the development of an organisation's evaluation or MEL policy (Trochim, 2009). ECB initiatives can thus be designed with the goal of institutionalising SAMEL within the organisation through the development of an institutionalised SAMEL policy.

If an ECB process precedes the SAMEL workshops, NPOs might be encouraged to direct more time and strategic resources towards MEL and integrate MEL more firmly across the organisation, thus 'spreading the load' of the MEL function. This would make capacity less

of a concern in implementing a systemic MEL framework. However, this hypothesis requires further research, particularly because it requires convincing NPOs possibly without a MEL culture to invest time and resources in MEL *and* ECB.

Finally, while I took a fairly conservative approach, a systems thinking approach to MEL can be argued to be a radical departure from traditional MEL methodologies. The SAMEL tests demonstrated that any transformation of MEL design necessitates a conducive and supportive context, especially with regard to programme funders. To properly benefit from the advantages of systems thinking, NPOs need to have the power and agency to enact the insights gained from a process such as SAMEL, particularly in intervention design.

Systems thinking is a powerful tool for critically reflecting on intervention design; however, much of the current non-profit funding landscape is incongruent with the core tenets of systems thinking. For example, developing collaborative relationships cannot take place in a funding landscape in which NPOs compete with each other. Attention cannot be paid to long-term, emergent outcome pathways when NPOs receive short-term grants that require short-term outcome reporting. Adaptive learning can only be embedded in an organisation that has been given the authority by funders to deviate from its original objectives. As such, “evaluators are constrained in their use of systems approaches by the evaluation circumstances in which they work” (Gates, 2017, p. 167). A transformation in MEL would thus require an associated transformation amongst the funding landscape in which MEL is used as a governance and accountability tool. While some global funding organisations are open to, and indeed encourage, a systems approach to programming, this conversation needs to include all funders, particularly those active in South Africa. One particular avenue worth exploring is participatory approaches that blend ‘downwards accountability’ — where the obligation is to report to programme participants — with upward accountability

(reporting to funders) (Urwin et al., 2023, p. 403). This is congruent with a Made in Africa Evaluation paradigm, which prioritises communities' perspectives of the intervention (Chilisa & Mertens, 2021).

This research has demonstrated that systems thinking is immensely valuable beyond wicked problems, and it is thus necessary to direct efforts towards mainstreaming systems thinking in MEL for all types of programmes in social development. The evaluation field, including the funders who commission MEL reporting and evaluations, requires a paradigm shift. Shifting paradigms and mental models are amongst the most powerful ways to create change in systems (Meadows, 2009). It is critical that evaluators recognise the power of systems thinking in all programmatic situations, not just those deemed complex. As Buckton et al. (2024) emphasise: evaluands must be understood as nested systems with complex, interdependent relationships and interactions with other systems. Relatedly, funders and evaluation commissioners must recognise that all social programmes involve these types of complex social change processes, requiring a conceptual lens that reflects this reality for effective programme design and MEL. As calls grow to transform evaluation to cope with the polycrisis, evaluators and commissioners must adopt frameworks that are driven by the principles of complexity (Buckton et al., 2024).

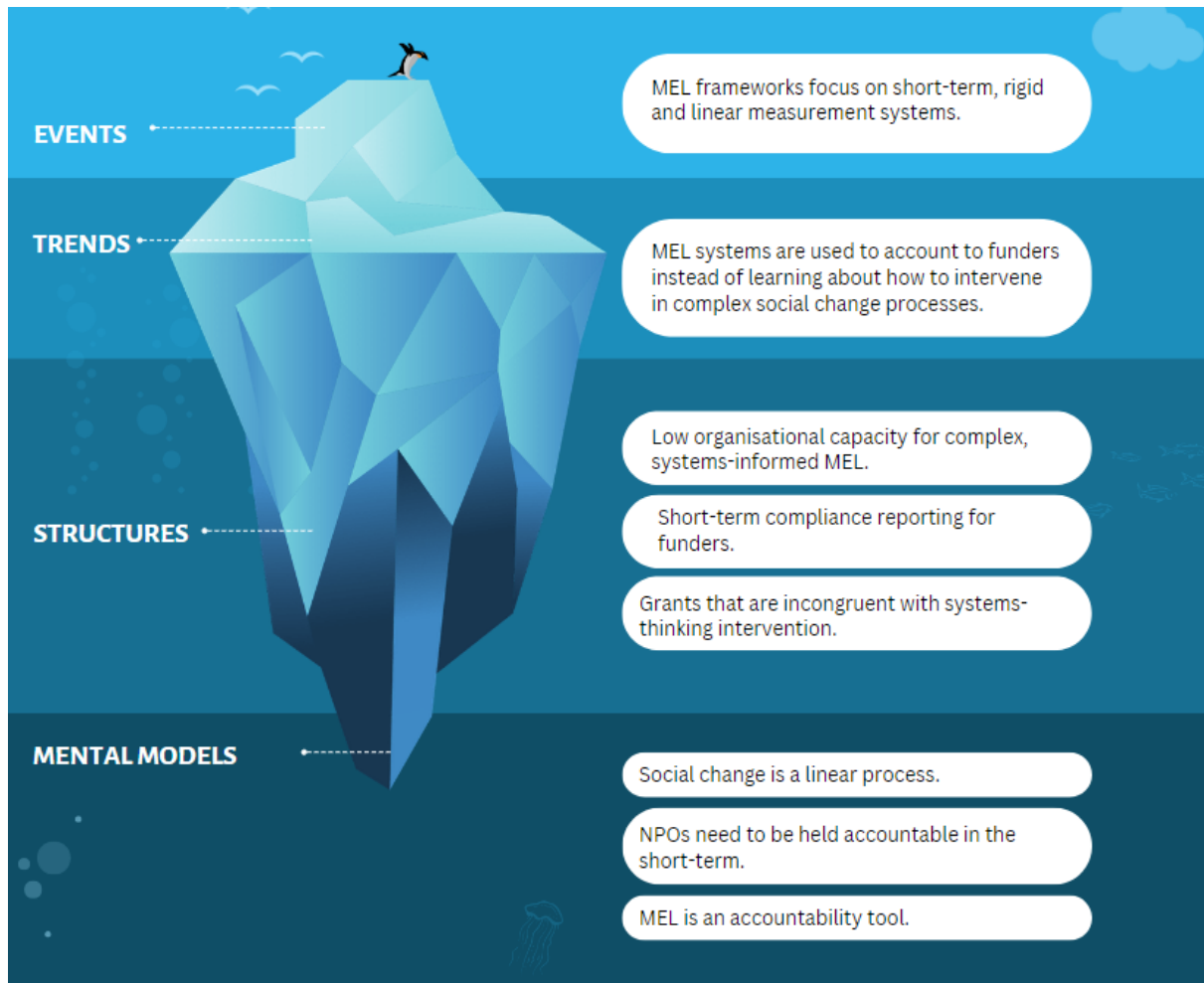
In Figure 54, I illustrate the systemic layers of the 'problem' of MEL and systems thinking¹¹. This study illustrates the structural barriers that prevent institutionalising a systems thinking approach to MEL in NPOs. As such, future research is needed to explore

¹¹ The Iceberg Model is often used to illustrate the underlying layers of a problem, highlighting what we can typically see (events) versus what is 'below the water' and generally ignored (structures and mental models) (Cunliff, 2004). Furthermore, when applied to a particular problem, it helps to differentiate between the problem's symptoms (events and patterns) and its root causes (structure and mental models) (Stroh, 2015).

ways in which these structural barriers could be negotiated, particularly in terms of the intersection of MEL and NPO accountability and governance.

Figure 54

Iceberg Model Summarising the Key Findings of this Research



Note. Adapted from "Connecting Systems Thinking and Action", by E. Cunliff, 2004, *Systems Thinking*, 15(2), p. 7 (<https://thesystemsthinker.com/wp-content/uploads/pdfs/150202pk.pdf>). Copyright 2004 by Pegasus Communications, Inc.

Limitations

This study, like any other that relies on purposive sampling, is at risk of selection bias. Additionally, as a qualitative study with a contextualist view of reality, the findings are tied to the context within which the data was collected and analysed (Madill et al., 2000). It is therefore unclear how transferable the study findings are to other NPOs or evaluators. Further, as the SAMEL Toolkit was tested only in the ECD sector, it is unclear whether the findings may be transferable to NPOs working in other social sectors. The NPOs that chose to participate in the workshops likely had inherent characteristics that influenced the SAMEL tests, such as interest and need for MEL, as well as an expressed need to review their programme strategies. Likewise, Heather's professional and academic background likely influenced how she engaged with the SAMEL Toolkit and facilitated the EYEN workshops. Given our similar backgrounds, it is thus unknown how other evaluators might have experienced the SAMEL Toolkit and its facilitation.

Further, this study's findings are limited by the small number of cases. Ideally, the SAMEL Toolkit would have been field-tested in several organisations, using several different evaluators, but this was not possible due to limited resources available to conduct the study. However, given the similarity in findings across the tests, it is not deemed a substantial limitation.

The nature of the tests themselves limited the focus group data collected during the workshops. Both NPOs regularly communicated their gratitude for their inclusion in the study, as MEL development is generally a significant expense for NPOs. It is, therefore, possible that the teams were hesitant to communicate dissatisfaction with the process or with the SAMEL content for fear of seeming ungrateful. While focus group moderators encouraged them to be honest and constructively critical, criticism was not forthcoming. It is unclear whether this

was because they had little critique to offer or whether they held back purposefully. While this limitation affects the focus group data, additional triangulated methods were employed to compensate for this (i.e., observations and evaluator accounts). Observational data did not indicate any dissatisfaction with the SAMEL process or particular content areas, outside of the confusion reported on in Chapter Eight. The teams' eagerness to work with Heather and me after the tests also indicates that the positive sentiments expressed in the focus groups are likely a true reflection of the teams' thoughts.

Finally, while described in the methodology section, my own bias is an important limitation to reiterate here. My multifaceted role as researcher, Toolkit developer, and test facilitator meant I facilitated the EduSpark test with a lot more systems thinking expertise than Heather did, and arguably more motivation for it to go well. Furthermore, my presence in the EYEN workshops as an observer arguably impacted the 'real-world' nature of the test. However, I believe that I managed my bias well by remaining highly reflexive during both the test- and analysis phases. I considered it important to remain honest about my own shortcomings and the Toolkit's limitations, as to ensure the research would be as meaningful and realistic as possible. The potential for bias was also weighed up against the potential advantages of being immersed in the research process. Being actively involved in all stages of the research, including both tests, gave me an incredibly in-depth and rich understanding of the data. Additionally, I was able to develop competence and expertise in all facets of the research process. I believe that the value of this involvement outweighed the potential risk of bias.

Contribution of New Knowledge

This exploratory study tested a systems thinking approach to theory-based MEL development amongst ECD NPOs, using a workshop format. The Toolkit's development

process was described, including the rationale for all tools and concepts. Expert feedback was obtained and incorporated, and the experience of facilitating the process was field-tested and reported on qualitatively. The contributions of this study are presented below.

Theoretical Contributions. To the best of my knowledge, the SAMEL Toolkit is the first set of guidelines that incorporates systems thinking tools and concepts into workshop-based, theory-based MEL development. While OM (Earl et al., 2001) is the closest comparison, it does not explicitly draw on the systems concepts approach, or teach evaluators about established systems thinking tools or existing methodologies. This study thus contributes a developed and tested knowledge product that, upon refinement, evaluators could use to support a MEL development process.

Methodological Contributions. Little background guidance has been made available on how systems-informed evaluation guidelines and products were developed, making such a process opaque and difficult to replicate. In this study, I openly described the experience of a newcomer's foray into systems thinking and MEL, documenting the rationale and process behind the SAMEL Toolkit's development, expert views on the content, and my and Heather's experience of facilitating workshops that engaged systems thinking in a MEL context. As such, this study serves as a much-needed test case into how evaluators could make sense of and apply systems thinking in their work.

Practical Contributions. As attention to the relationship between systems thinking and evaluation grows, it is imperative to understand the real-world advantages of blending the fields, as well as the barriers to doing so effectively. This research responds to an identified gap in the literature on how evaluation stakeholders, including evaluation commissioners or clients, experience the use of systems thinking (Gates, 2017; Walton, 2016).

As systems thinking continues to receive growing support in evaluation, it is critical to test innovative ways in which these disciplines can benefit the field in ways other than those already documented (i.e., summative evaluation, developmental evaluation, and adaptive monitoring in complex programmes). This includes understanding how evaluators (emerging and experienced) can engage with systems thinking in real-world contexts, and whether systems thinking is advantageous in fairly non-complex settings such as small NPOs delivering programmes with 'known' or probable outcomes.

This study has demonstrated the impactful potential of systems thinking for MEL development in non-profit settings, above and beyond 'wicked problems'. By facilitating engagement with sectoral systems, relationships with other actors, partnerships, programme boundaries and the organisation's context – the SAMEL Toolkit tests demonstrated that NPO teams can benefit from enhanced strategic planning and programme design (by way of programme theory). Additionally, this research emphasises the critical role of ECB, and more flexible funding and reporting structures, in empowering organisations to implement systems thinking effectively in MEL. By bridging a sizeable gap in the literature, this study contributes an approach that supports NPOs to be better equipped to navigate, and respond to, the complexities inherent to social change, thereby supporting a growing movement toward the transformation of evaluation. The SAMEL Toolkit thus contributes to a future in which evaluation is evolved to more effectively support the resilience and adaptation needed in social development today.

References

- Abrahams, M. A. (2015). A review of the growth of monitoring and evaluation in South Africa: Monitoring and evaluation as a profession, an industry and a governance tool. *African Evaluation Journal*, 3(1), Article 142. <https://doi.org/10.4102/aej.v3i1.142>
- Ackoff, R. L. (1981). *Creating the corporate future*. Wiley.
- Adams, J., & Garbutt, A. (2008). *Participatory monitoring and evaluation in Practice: Lessons learnt from Central Asia*. INTRAC. <https://www.intrac.org/wpcms/wp-content/uploads/2016/09/Praxis-Paper-21-PME-in-Practice-Jerry-Adams-and-Anne-Garbutt.pdf>
- African Evaluation Association (AfrEA). (2007). *African Evaluation Guidelines - Standards and norms*. African Evaluation Association. https://member.samea.org.za/img/AEG_2007_EN.pdf
- African Evaluation Association (AfrEA). (2021). *The African evaluation principles: 2021 version*. <https://afrea.org/AEP/new/The-African-Evaluation-Principles.pdf>
- Agazzi, E. (1978). Systems theory and the problem of reductionism. *Erkenntnis*, 12(3), 339–358. <https://doi.org/10.1007/BF00216104>
- Apgar, J. M., Allen, W., Albert, J., Douthwaite, B., Paz Ybarnegaray, R., & Lunda, J. (2017). Getting beneath the surface in program planning, monitoring and evaluation: Learning from use of participatory action research and theory of change in the CGIAR Research Program on Aquatic Agricultural Systems. *Action Research*, 15(1), 15–34. <https://doi.org/10.1177/1476750316673879>
- Arkesteijn, M., Van Mierlo, B., & Leeuwis, C. (2015). The need for reflexive evaluation approaches in development cooperation. *Evaluation*, 21(1), 99–115. <https://doi.org/10.1177/1356389014564719>
- Arnold, D. H., Kupersmidt, J. B., Voegler-Lee, M. E., & Marshall, N. A. (2012). The association between preschool children's social functioning and their emergent academic skills. *Early Childhood Research Quarterly*, 27(3), 376–386. <https://doi.org/10.1016/j.ecresq.2011.12.009>
- Arnold, R. D., & Wade, J. P. (2015). A Definition of Systems Thinking: A Systems Approach. *Procedia Computer Science*, 44, 669–678. <https://doi.org/10.1016/j.procs.2015.03.050>
- Ashby, W. R. (1956). *Introduction to cybernetics*. Wiley.
- Ashley-Cooper, M., van Niekerk, L.-J., & Atmore, E. (2019). Early Childhood Development in South Africa: Inequality and Opportunity. In N. Spaull & J. D. Jansen (Eds.), *South*

- African Schooling: The Enigma of Inequality* (pp. 87–108). Springer International Publishing. https://doi.org/10.1007/978-3-030-18811-5_5
- Atmore, E. (2012). Challenges facing the early childhood development sector in South Africa. *South African Journal of Childhood Education*, 2(1), 120–139. <https://doi.org/10.4102/sajce.v2i1.25>
- Atmore, E. (2013). Early childhood development in South Africa – progress since the end of apartheid. *International Journal of Early Years Education*, 21(2/3), 152–162. <https://doi.org/10.1080/09669760.2013.832941>
- Attenborough, K. (2006). Soft systems in a hardening world: Evaluating urban regeneration. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation: An expert anthology* (pp. 75–87). American Evaluation Association.
- Barbrook-Johnson, P., Castellani, B., Hills, D., Penn, A., & Gilbert, N. (2021). Policy evaluation for a complex world: Practical methods and reflections from the UK Centre for the Evaluation of Complexity across the Nexus. *Evaluation*, 27(1), 4–17. <https://doi.org/10.1177/1356389020976491>
- Bayat, A., & Madyibi, S. (2022). Impact of COVID-19 on South Africa’s early childhood education and the Sustainable Development Goals. In W. Leal Filho, I. R. Abubakar, I. da Silva, R. Pretorius, & K. Tarabieh (Eds.), *SDGs in Africa and the Middle East Region* (pp. 1–27). Springer International Publishing. https://doi.org/10.1007/978-3-030-91260-4_81-1
- Beer, S. (1972). *Brain of the firm: A development in management cybernetics*. Herder & Herder.
- Befani, B., Ramalingam, B., & Stern, E. (2015). Introduction—Towards systemic approaches to evaluation and impact. *IDS Bulletin*, 46(1), 1–6. <https://doi.org/10.1111/1759-5436.12116>
- Bem-Haja, P., Nossa, P., Pereira, D. S., & Silva, C. F. (2022). Did the COVID-19 pandemic lockdown harm pre-schoolers learning in Portugal? Yes, but with variations depending on socio-economic status. *Education Sciences*, 12(10), 710. <https://doi.org/10.3390/educsci12100710>
- Beney, T., Mathe, J., Ntakumba, S., Basson, R., Naidu, V., & Leslie, M. (2015). A reflection on the partnership between government and South African Monitoring and Evaluation Association. *African Evaluation Journal*, 3(1), Article 164. <https://doi.org/10.4102/aej.v3i1.164>
- Bevan, G., & Hood, C. (2006). What’s measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–538. <https://doi.org/10.1111/j.1467-9299.2006.00600.x>

- Biersteker, L., Dawes, A., Hendricks, L., & Tredoux, C. (2016). Center-based early childhood care and education program quality: A South African study. *Early Childhood Research Quarterly*, 36, 334–344. <https://doi.org/10.1016/j.ecresq.2016.01.004>
- Blaser Mapitsa, C., & Khumalo, L. (2018). Diagnosing monitoring and evaluation capacity in Africa. *African Evaluation Journal*, 6(1), Article a255. <https://doi.org/10.4102/aej.v6i1.255>
- Bornstein, L. (2003). Management standards and development practice in the South African aid chain. *Public Administration and Development*, 23(5), 393–404. <https://doi.org/10.1002/pad.291>
- Bornstein, L. (2006). Systems of Accountability, webs of deceit? Monitoring and evaluation in South African NGOs. *Development*, 49(2), 52–61. <https://doi.org/10.1057/palgrave.development.1100261>
- Bourgeois, I., Hart, R. E., Townsend, S. H., & Gagné, M. (2011). Using hybrid models to support the development of organizational evaluation capacity: A case narrative. *Evaluation and Program Planning*, 34(3), 228–235. <https://doi.org/10.1016/j.evalprogplan.2011.03.003>
- Bowman, N. (2020). Nation-to-nation in evaluation: Utilizing an indigenous evaluation model to frame systems and government evaluations. *New Directions for Evaluation*, 2020(166), 101 – 118. <https://doi.org/10.1002/ev.20411>
- Boyd, A., Brown, M., & Midgley, G. (2004). Systemic intervention for community OR: Developing services with young people (under 16) living on the streets. In G. Midgley & A. Ochoa-Arias (Eds.), *Community Operational Research: OR and Systems Thinking for Community Development* (pp. 203–252). Springer US.
- Boyden, J., Dawes, A., Dornan, P., & Tredoux, C. (2019). *Tracing the consequences of child poverty: Evidence from the Young Lives study in Ethiopia, India, Peru and Vietnam*. Policy Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage Publications.
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. Sage Publications.
- Brewer, J., & Hunter, A. (2006). *Foundations of multimethod research: Synthesizing styles*. Sage Publications.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

- Brown, A., & Green, T. D. (2016). *The essentials of instructional design: Connecting fundamental principles with process and practice* (2nd ed.). Routledge.
- Buckton, S. J., Fazey, I., Ball, P., Ofir, Z., Colvin, J., Darby, M., Hejnowicz, A. P., Leicester, G., Newman, R., Page, G., Parsons, K., & van Mierlo, B. (2024). Twelve principles for transformation-focused evaluation. SSRN. <https://doi.org/10.2139/ssrn.4872604>
- Buetti, D., Bourgeois, I., & Jafary, M. (2023). Examining the competencies required by evaluation capacity builders in community-based organizations. *Evaluation and Program Planning*, *97*, 102242. <https://doi.org/10.1016/j.evalprogplan.2023.102242>
- Burke, D. D. (2006). Systems dynamics-based computer simulations and evaluation. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation: An expert anthology* (pp. 47–60). American Evaluation Association.
- Burnes, B. (2007). Kurt Lewin and the Harwood Studies: The foundations of OD. *The Journal of Applied Behavioral Science*, *43*(2), 213–231. <https://doi.org/10.1177/0021886306297004>
- Burnes, B., & Cooke, B. (2013). Kurt Lewin's Field Theory: A review and re-evaluation. *International Journal of Management Reviews*, *15*, 408–425. <https://doi.org/10.1111/j.1468-2370.2012.00348.x>
- Bustamante, M., Vidueira, P., & Baker, L. (2021). Systems thinking and complexity science-informed evaluation frameworks: Assessment of the economics of ecosystems and biodiversity for agriculture and food. *New Directions for Evaluation*, *2021*(170), 81–100. <https://doi.org/10.1002/ev.20455>
- Cabrera, D., Cabrera, L., & Midgley, G. (2023). The Four Waves of Systems Thinking. *Journal of Systems Thinking*, *3*, 1–51. <https://doi.org/10.54120/jost.000051>
- Cabrera, D., Cabrera, L., & Powers, E. (2015). A unifying theory of systems thinking with psychosocial applications. *Systems Research and Behavioral Science*, *32*(5), 534–545. <https://doi.org/10.1002/sres.2351>
- Cabrera, D., & Colosi, L. (2008). Distinctions, systems, relationships, and perspectives (DSRP): A theory of thinking and of things. *Evaluation and Program Planning*, *31*(3), 311–317. <https://doi.org/10.1016/j.evalprogplan.2008.04.001>
- Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning*, *31*(3), 299–310. <https://doi.org/10.1016/j.evalprogplan.2007.12.001>
- Carman, J. G. (2011). Understanding evaluation in nonprofit organizations. *Public Performance & Management Review*, *34*(3), 350–377. <https://doi.org/10.2753/PMR1530-9576340302>

- Chapman, S., Boodhoo, A., Duffy, C., Goodman, S., & Michalopoulou, M. (2023). Theory of change in complex research for development programmes: Challenges and solutions from the global challenges research fund. *The European Journal of Development Research*, 35(2), 298–322. <https://doi.org/10.1057/s41287-023-00574-0>
- Chaudhary, A. K., Diaz, J., Jayaratne, K. S. U., & Assan, E. (2020). Evaluation capacity building in the nonformal education context: Challenges and strategies. *Evaluation and Program Planning*, 79, 101768. <https://doi.org/10.1016/j.evalprogplan.2019.101768>
- Checkland, P. (1981). *Systems thinking, systems practice*. Wiley.
- Checkland, P., & Poulter, J. (2006). *Learning for action: A short definitive account of soft systems methodology and its use for practitioner, teachers, and students*. Wiley.
- Checkland, P., & Poulter, J. (2010). Soft systems methodology. In M. Reynolds & S. Holwell (Eds.), *Systems approaches to managing change: A practical guide* (pp. 191–242). Springer.
- Chen, H. (2015). *Practical program evaluation: Theory-driven evaluation and the integrated evaluation perspective* (2nd ed.). Sage Publications.
- Chen, H.-T., & Rossi, P. H. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review*, 7(3), 283–302. <https://doi.org/10.1177/0193841X8300700301>
- Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. *Educational Researcher*, 21(6), 13–17. <https://doi.org/10.3102/0013189X021006013>
- Cherryholmes, C. H. (1994). More notes on pragmatism. *Educational Researcher*, 23(1), 16–18. <https://doi.org/10.2307/1176282>
- Chilisa, B., & Mertens, D. M. (2021). Indigenous Made in Africa Evaluation frameworks: Addressing epistemic violence and contributing to social transformation. *American Journal of Evaluation*, 42(2), 241–253. <https://doi.org/10.1177/1098214020948601>
- Churchman, C. W. (1970). Operations research as a profession. *Management Science*, 17, 37–53. <https://doi.org/10.1287/mnsc.17.2.B37>
- Clinton, J. M., & Hattie, J. (2021). Cognitive complexity of evaluator competencies. *Evaluation and Program Planning*, 89, 102006. <https://doi.org/10.1016/j.evalprogplan.2021.102006>
- Collie, R. J., Martin, A. J., Nassar, N., & Roberts, C. L. (2019). Social and emotional behavioral profiles in kindergarten: A population-based latent profile analysis of links to socio-educational characteristics and later achievement. *Journal of Educational Psychology*, 111(1), 170–187. <https://doi.org/10.1037/edu0000262>

- Connor, C. M., Son, S.-H., Hindman, A. H., & Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and early reading outcomes. *Journal of School Psychology, 43*(4), 343–375. <https://doi.org/10.1016/j.jsp.2005.06.001>
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). Sage Publications.
- Coryn, C. L. S. (2005). Book review: Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness. *American Journal of Evaluation, 26*(3), 405–407. <https://doi.org/10.1177/1098214005278767>
- Coryn, C. L. S. (2008). Book review: Program Theory-Driven Evaluation Science: Strategies and Applications, by Stewart I. Donaldson. Mahwah, NJ: Lawrence Erlbaum, 2007. *American Journal of Evaluation, 29*(2), 215–217. <https://doi.org/10.1177/1098214008314981>
- Coryn, C. L. S., Noakes, L. A., Westine, C. D., & Schröter, D. C. (2011). A systematic review of theory-driven evaluation practice from 1990 to 2009. *American Journal of Evaluation, 32*(2), 199–226. <https://doi.org/10.1177/1098214010389321>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Sage Publications.
- Dahler-Larsen, P. (2018). Theory-based evaluation meets ambiguity: The role of Janus variables. *American Journal of Evaluation, 39*(1), 6–23. <https://doi.org/10.1177/1098214017716325>
- Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science, 10*(1), 49. <https://doi.org/10.1186/s13012-015-0237-x>
- Davies, R. (2004a). Scale, complexity and the representation of Theories of Change. *Evaluation, 10*(1), 101–121. <https://doi.org/10.1177/1356389004043124>
- Davies, R. (2004b). Scale, complexity and the representation of Theories of Change. *Evaluation, 10*(1), 101–121. <https://doi.org/10.1177/1356389004043124>
- Dawes, A., Biersteker, L., Snelling, M., & Horler, J. (2020). *The early learning outcomes study: Technical report*. Innovation Edge. <https://innovationedge.org.za/wp-content/uploads/2020/05/IE-ELPO-Study-Technical-Report-FINAL.pdf>
- Dawes, A., Biersteker, L., Snelling, M., Horler, J., & Girdwood, E. (2021). To what extent can community-based playgroup programmes targeting low-income children improve learning outcomes prior to entering the reception year in south africa? A quasi-

experimental field study. *Early Education and Development*, 34(1), 256–273.
<https://doi.org/10.1080/10409289.2021.2005748>

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *The SAGE PUBLICATIONS handbook of qualitative research* (3rd ed.). Sage Publications.

Department of Basic Education. (2022). *ECD Census 2021 Report*. Department of Basic Education. <https://datadrive2030.co.za/wp-content/uploads/2022/09/ecdc-2021-report.pdf>

Desmond, C., Viviers, A., Edwards, T., Rich, K., Martin, P., & Richter, L. (2019). Priority-setting in the roll out of South Africa's National Integrated ECD Policy. *Early Years*, 39(3), 276–294. <https://doi.org/10.1080/09575146.2019.1572074>

Dlakavu, A., Mathebula, J., & Mkhize, S. (2022). Decolonising and indigenising evaluation practice in Africa: Roadmap for mainstreaming the Made in Africa Evaluation approach. *African Evaluation Journal*, 10(1), a620.
<https://doi.org/10.4102/aej.v10i1.620>

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications*. Lawrence Erlbaum Associates.

Donaldson, S. I. (2021). *Introduction to theory-driven program evaluation: Culturally responsive and strengths-focused applications* (2nd ed.). Routledge.

Douthwaite, B., Ahmad, F., & Shah, G.-M. (2020). Putting theory of change into use in complex settings. *Canadian Journal of Program Evaluation*, 35(1), 35–52.
<https://doi.org/10.3138/cjpe.43168>

Douthwaite, B., & Hoffecker, E. (2017). Towards a complexity-aware theory of change for participatory research programs working within agricultural innovation systems. *Agricultural Systems*, 155, 88–102. <https://doi.org/10.1016/j.agsy.2017.04.002>

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>

Duncan, G. J., Magnuson, K., Kalil, A., & Ziol-Guest, K. (2011). The importance of early childhood poverty. *Social Indicators Research*, 108(1), 87–98.
<https://doi.org/10.1007/s11205-011-9867-9>

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41(3–4), 327–350.
<https://doi.org/10.1007/s10464-008-9165-0>

- Dusenbury, L. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237–256. <https://doi.org/10.1093/her/18.2.237>
- Dyehouse, M., Bennett, D., Harbor, J., Childress, A., & Dark, M. (2009). A comparison of linear and systems thinking approaches for program evaluation illustrated using the Indiana Interdisciplinary GK-12. *Evaluation and Program Planning, 32*(3), 187–196. <https://doi.org/10.1016/j.evalprogplan.2009.03.001>
- Earl, S., Carden, F., Patton, M. Q., & Smutylo, T. (2001). *Outcome mapping: Building learning and reflection into development programs*. International Development Research Centre.
- Ebrahim, H. B., Martin, C., & Excell, L. (2021). Early childhood teachers' and managers' lived experiences of the COVID-19 pandemic in South Africa. *Journal of Education, 84*, 1–18. <https://doi.org/10.17159/2520-9868/i84a11>
- Emery, F. E. (Ed.). (1969). *Systems thinking*. Penguin Press.
- Eoyang, G. H. (2006). Human systems dynamics: Complexity-based approach to a complex evaluation. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation: An expert anthology* (pp. 123–140). American Evaluation Association.
- Essel, E. A. (2021). The demand and effects of accountability and efficient utilization of resources by donors of local nongovernmental organisations (NGOs): A qualitative assessment of some selected NGOs in the central region of Ghana. *European Scientific Journal, ESJ, 17*(32), 21. <https://doi.org/10.19044/esj.2021.v17n32p21>
- Eyben, R. (2013). *Uncovering the politics of 'evidence' and 'results'. A framing paper for development practitioners*. Big Push Forward. <https://bigpushforward.net/wp-content/uploads/2011/01/The-politics-of-evidence-11-April-20132.pdf>
- Faust, J., & Verspohl, I. (2018). *The OECD-DAC evaluation criteria: To reform or transform? DEval policy brief 9/2018*. German Institute for Development Evaluation. https://www.deval.org/fileadmin/Redaktion/PDF/05-Publikationen/Policy_Briefs/2018_9_DAC_Evaluierungskriterien/DEval_Policy_Brief_DAC_Evaluation_Criteria_2018_EN.pdf
- Field, E., Vila, M., Runk, L., Mactaggart, F., Rosewell, A., & Nathan, S. (2018). Lessons for health program monitoring and evaluation in a low resource setting. *Rural and Remote Health, 18*(4), 1–9. <https://doi.org/10.22605/RRH4596>
- Fitzpatrick, C., McKinnon, R. D., Blair, C. B., & Willoughby, M. T. (2014). Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learning and Instruction, 30*, 25–31. <https://doi.org/10.1016/j.learninstruc.2013.11.003>

- Flood, R. C., & Jackson, M. C. (1991). *Critical systems thinking: Directed readings*. Wiley.
- Flood, R. L. (Ed.). (1997). *Critical systems thinking: Current research and practice*. Plenum Press.
- Forrester, J. W. (2007a). System dynamics—A personal view of the first fifty years. *System Dynamics Review*, 23(2–3), 345–358. <https://doi.org/10.1002/sdr.382>
- Forrester, J. W. (2007b). System dynamics—The next fifty years. *System Dynamics Review*, 23(2– 3), 359–370. <https://doi.org/10.1002/sdr.381>
- Fredericks, K. A., Deegan, M., & Carman, J. G. (2008). Using system dynamics as an evaluation tool: Experience from a demonstration program. *American Journal of Evaluation*, 29(3), 251–267. <https://doi.org/10.1177/1098214008319446>
- French, M., Hesselgreaves, H., Wilson, R., Hawkins, M., & Lowe, T. (2023). *Harnessing complexity for better outcomes in public and non-profit services*. Policy Press.
- Funnel, S. C., & Rogers, P. J. (2011). *Purposeful program theory*. Jossey-Bass.
- Gates, E. F. (2016). Making sense of the emerging conversation in evaluation about systems thinking and complexity science. *Evaluation and Program Planning*, 59, 62–73. <https://doi.org/10.1016/j.evalprogplan.2016.08.004>
- Gates, E. F. (2017). Learning from seasoned evaluators: Implications of systems approaches for evaluation practice. *Evaluation*, 23(2), 152–171. <https://doi.org/10.1177/1356389017697613>
- Gates, E. F. (2018). Toward valuing with Critical Systems Heuristics. *American Journal of Evaluation*, 39(2), 201–220. <https://doi.org/10.1177/1098214017703703>
- Gates, E. F., Walton, M., Vidueira, P., & McNall, M. (2021). Introducing systems- and complexity-informed evaluation. *New Directions for Evaluation*, 2021(170), 13–25. <https://doi.org/10.1002/ev.20466>
- Georgiadis, A., & Penny, M. E. (2017). Child undernutrition: Opportunities beyond the first 1000 days. *The Lancet Public Health*, 2(9), e399. [https://doi.org/10.1016/S2468-2667\(17\)30154-8](https://doi.org/10.1016/S2468-2667(17)30154-8)
- Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, 21(2), 135–146. <https://doi.org/10.1057/ejis.2011.54>
- Goncharenko, G. (2021). The multiplicity of logics, trust, and interdependence in donor-imposed reporting practices in the nonprofit sector. *Financial Accountability & Management*, 37(2), 124–141. <https://doi.org/10.1111/faam.12241>

- González, M., Loose, T., Liz, M., Pérez, M., Rodríguez-Vinçon, J. I., Tomás-Llerena, C., & Vásquez-Echeverría, A. (2022). School readiness losses during the COVID-19 outbreak. A comparison of two cohorts of young children. *Child Development, 93*(4), 910–924. <https://doi.org/10.1111/cdev.13738>
- Gopalakrishnan, S., & Clarke, T. (2015). *System mapping: A guide to developing actor maps*. Foundation Strategy Group. <http://mcbc.info/wp-content/uploads/2016/06/Final-System-Mapping-Guide.pdf>
- Goremucheche, R. (2022, March 3). *Patience and humility: What we often overlook when preparing to facilitate a workshop*. Wits University. <https://www.wits.ac.za/clear-aa/articles/patience-and-humility-what-we-often-overlook-when-preparing-to-facilitate-a-workshop.html>
- Görgens-Albino, M., & Kusek, J. Z. (2009). *Making monitoring and evaluation systems work: A capacity development toolkit*. World Bank.
- Greene, J. C. (2007). *Mixed methods in social inquiry* (1st ed.). Jossey-Bass.
- Gregory, A. J., & Jackson, M. C. (1992). Evaluating organizations: A systems and contingency approach. *Systems Practice, 5*(1), 37–60. <https://doi.org/10.1007/BF01060046>
- Gruening, G. (2001). Origin and theoretical basis of new public management. *International Public Management Journal, 4*(1), 1–25. [https://doi.org/10.1016/S1096-7494\(01\)00041-1](https://doi.org/10.1016/S1096-7494(01)00041-1)
- Hall, K., Sambu, W., Almeleh, C., Mabaso, K., Giese, S., & Proudlock, P. (2019). *South African early childhood review 2019*. Children’s Institute, University of Cape Town, and Ilifa Labantwana. https://ilifalabantwana.co.za/wp-content/uploads/2019/09/SA-ECR_2019_12_09_2019_online_pages.pdf
- Hamadani, J. D., Tofail, F., Hilaly, A., Huda, S. N., Engle, P., & Grantham-McGregor, S. M. (2010). Use of family care indicators and their relationship with child development in Bangladesh. *Journal of Health, Population and Nutrition, 28*(1), 23–33. <https://doi.org/10.3329/jhpn.v28i1.4520>
- Hargreaves, M. B. (2010). *Evaluating system change: A planning guide*. Mathematica Policy Research. <https://www.mathematica.org/our-publications-and-findings/publications/evaluating-system-change-a-planning-guide>
- Harnar, M. A., Hillman, J. A., Endres, C. L., & Snow, J. Z. (2020). Internal formative meta-evaluation: Assuring quality in evaluation practice. *American Journal of Evaluation, 41*(4), 603–613. <https://doi.org/10.1177/1098214020924471>
- Hassmiller Lich, K., Urban, J. B., Frerichs, L., & Dave, G. (2017). Extending systems thinking in planning and evaluation using group concept mapping and system dynamics to tackle

- complex problems. *Evaluation and Program Planning*, 60, 254–264.
<https://doi.org/10.1016/j.evalprogplan.2016.10.008>
- Hayes, N., O’Toole, L., & Halpenny, A. (2017). *Introducing Bronfenbrenner: A guide for practitioners and students in early years education*. Routledge.
- Heiss, A., & Kelley, J. G. (2017). Between a rock and a hard place: International NGOs and the dual pressures of donors and host governments. *The Journal of Politics*, 79(2), 732 – 741. <http://dx.doi.org/10.1086/691218>
- Hevner, A., & Chatterjee, S. (2010). Design Science Research in Information Systems. In A. Hevner & S. Chatterjee, *Design Research in Information Systems* (Vol. 22, pp. 9–22). Springer US. https://doi.org/10.1007/978-1-4419-5653-8_2
- Holland, J. H. (1992). Complex adaptive systems. *Daedalus*, 121(1), 17-30.
- Holling, C. S., & Sundstrom, S. M. (2015). Adaptive management, a personal history. In C. R. Allen & A. S. Garmestani (Eds.), *Adaptive Management of Social-Ecological Systems* (1st ed.). Springer. <https://doi.org/10.1007/978-94-017-9682-8>
- Horler, J. (2021). *Early learning in South Africa: A funder survey*. IPASA, Tutuwa Community Foundation, Innovation Edge. <https://ipa-sa.org.za/download/early-learning-sa-funder-survey-report-december-2021/?wpdmdl=4862&refresh=660e852f4d0741712227631>
- Horler, J., & Biersteker, L. (2022). *ECD funding landscape: Survey*. IPASA, Tutuwa Community Foundation. <https://ipa-sa.org.za/about/funder-support-initiatives/education-funder-support-initiatives/ipasa-ecd-funding-landscape-survey-october-2022/>
- Horler, J., Biersteker, L., & Berry, L. (2023). *Survey of resource and training organisations supporting early learning programmes in South Africa*. National ECD Alliance and Umncedi. <https://www.ecdalliance.org/wp-content/uploads/2023/09/Resource-and-Training-Organisations-Skills-Survey-2023-Report.pdf>
- Horler, J., Snelling, M., Biersteker, L., Dawes, A., & Girdwood, E. (2019). *SmartStart outcome evaluation: Final report* (p. 211). SmartStart. <https://www.smartstart.org.za/wp-content/uploads/2020/05/Smartstart-Outcome-Evaluation-FINAL-FULL-TECHNICAL-REPORT.pdf>
- Hsiao, C., & Richter, L. M. (2014). Early mental development as a predictor of preschool cognitive and behavioral development in South Africa: The moderating role of maternal education in the birth to twenty cohort. *Infants & Young Children*, 27(1), 74–87. <https://doi.org/10.1097/IYC.0000000000000002>
- Hummelbrunner, R. (2011). Systems thinking and evaluation. *Evaluation*, 17(4), 395–403. <https://doi.org/10.1177/1356389011421935>

- Hunter, A., & Brewer, J. (2015). Designing multimethod research. In S. Hesse-Biber & R. B. Johnson (Eds.), *The Oxford handbook of multimethod and mixed methods research inquiry* (pp. 185–205). Oxford University Press.
- Ilifa Labantwana. (2014). *The essential package*. Ilifa Labantwana. <https://ilifalabantwana.co.za/wp-content/uploads/2017/06/The-Essential-Package-2ed.pdf>
- Jackson, M. C. (1982). The nature of soft systems thinking: The work of Churchman, Ackoff and Checkland. *Journal of Applied Systems Analysis*, 9, 17–29.
- Jackson, M. C. (1987). New directions in management science. In M. C. Jackson & P. Keys (Eds.), *New directions in management science* (pp. 133–165). Gower.
- Jackson, M. C. (2008). *Systems thinking: Creative holism for managers*. Wiley.
- Jackson, M. C. (2020). Critical systems practice 1: Explore —Starting a multimethodological intervention. *Systems Research and Behavioral Science*, 37(5), 839–858. <https://doi.org/10.1002/sres.2746>
- Jackson, M. C. (2021). Critical systems practice 2: Produce —Constructing a multimethodological intervention strategy. *Systems Research and Behavioral Science*, 38(5), 594–609. <https://doi.org/10.1002/sres.2809>
- Jackson, M. C. (2022). Critical systems practice 3: Intervene —Flexibly executing a multimethodological intervention. *Systems Research and Behavioral Science*, 39(6), 1014–1023. <https://doi.org/10.1002/sres.2909>
- Jackson, M. C. (2023a). Critical systems practice 4: Check —Evaluating and reflecting on a multimethodological intervention. *Systems Research and Behavioral Science*, 40(4), 617–632. <https://doi.org/10.1002/sres.2912>
- Jackson, M. C. (2023b). In search of a golden mean for systemic evaluation: A response to Michael Quinn Patton. *Systems Research and Behavioral Science*, 40(4), 636–638. <https://doi.org/10.1002/sres.2938>
- Jacob, S. (2023). The institutionalization of evaluation around the globe: Understanding the main drivers and effects over the past decades. In R. Stockmann, W. Meyer, & L. Taube (Eds.), *Handbook of public policy evaluation* (pp. 187–205). Edward Elgar Publishing.
- Jennings, W., Stoker, G., Bunting, H., Valgarðsson, V. O., Gaskell, J., Devine, D., McKay, L., & Mills, M. C. (2021). Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy. *Vaccines*, 9(6), 593. <https://doi.org/10.3390/vaccines9060593>
- Jones, H. (2006). *Making Outcome Mapping work: Evolving experiences from around the world*. Outcome Mapping Learning Community. <https://idl-bnc->

idrc.dspacedirect.org/server/api/core/bitstreams/2846b9b4-104d-4b66-8064-932e84f3a1aa/content

- Kachur, D., Soal, S., & Van Blerk, R. (2016). Stretching between learning and accountability: Experiences of South African non-governmental organisations. *African Evaluation Journal*, 4(1), 11 pages. <https://doi.org/10.4102/aej.v4i1.71>
- Kania, J., Kramer, M., & Russell, P. (2014). *Strategic philanthropy for a complex world*. Stanford Social Innovation Review. http://www.oshkoshareacf.org/wp-content/uploads/2017/11/Strategic_Philanthropy_for_Complex_World-1.pdf
- Kehrer, D., Flossman-Kraus, U., Alarcon, S. V. R., Albers, V., & Aschmann, G. (2020). *Transforming our work: Getting ready for transformational projects*. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). <https://www.giz.de/expertise/html/61603.html>
- Kekana, M. J., Mabapa, N. S., & Mbhenyane, X. G. (2020). Impact of the child support grant on the diet and nutritional status of children under 5 years old in Mogalakwena Municipality, Limpopo Province, South Africa. *South African Journal of Child Health*, 14(1), 15–19. <https://doi.org/>. <https://doi.org/10.7196.SAJCH.2020.v14.i1.1598>
- Kelly, L. M. (2021). A clash of values: Deep-rooted discord between empowering, participatory, community-driven development and results-focused, evidence-based evaluation. *Community Development*, 52(5), 607–623. <https://doi.org/10.1080/15575330.2021.1936101>
- Kincheloe, J. L. (2001). Describing the bricolage: Conceptualizing a new rigor in qualitative research. *Qualitative Inquiry*, 7(6), 679–692.
- Kincheloe, J. L., & Berry, K. S. (Eds.). (2004). *Rigour and complexity in educational research: Conceptualizing the bricolage*. Open University Press.
- Knowlton, L. W., & Phillips, C. C. (2013). *The logic model guidebook: Better strategies for great results* (2nd ed.). Sage Publications.
- Kumaran, M., Samuel, N., & Winston, L. (2012). The NGO sector in South Africa: History, issues and prospects. *Journal for Development and Leadership*, 1(2), 31–45.
- Kusek, J. Z., & Rist, R. C. (2004). *Ten steps to a results-based monitoring and evaluation system: A handbook for development practitioners*. World Bank.
- Kushner, J., Cabrera, L., & Cabrera, D. (2021). Systems Evaluation (SysEval): Applying systems thinking to evaluation. In D. Cabrera, L. Cabrera, & G. Midgley (Eds.), *Routledge Handbook of Systems Thinking*. Routledge.

- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation, 33*(3), 307–338. <https://doi.org/10.1177/1098214011434608>
- Latham, N. (2014). *A practical guide to evaluating systems change in a human services system context*. Center for Evaluation Innovation. <https://www.evaluationinnovation.org/publication/a-practical-guide-to-evaluating-systems-change-in-a-human-services-system-context/>
- LeFevre, A. E., Dane, P., Copley, C. J., Pienaar, C., Parsons, A. N., Engelhard, M., Woods, D., Bekker, M., Benjamin, P., Pillay, Y., Barron, P., Seebregts, C. J., & Mohan, D. (2018). Unpacking the performance of a mobile health information messaging program for mothers (MomConnect) in South Africa: *Evidence on program reach and messaging exposure*. *BMJ Global Health, 3*(Suppl 2), e000583. <https://doi.org/10.1136/bmjgh-2017-000583>
- Lennie, J., & Tacchi, J. (2013). *Evaluating communication for development: A framework for social change*. Routledge.
- Lévi-Strauss, C. (1966). *The savage mind*. University of Chicago Press.
- Lincoln, Y. S., Lynham, S. A., & Guba, E. C. (2018). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage Publications handbook of qualitative research* (5th ed., pp. 222–264). Sage Publications.
- Lowe, T., & Wilson, R. (2017). Playing the game of outcomes-based performance management. Is gamesmanship inevitable? Evidence from theory and practice. *Social Policy & Administration, 51*(7), 981–1001. <https://doi.org/10.1111/spol.12205>
- Luo, C., Chen, A., Cui, B., & Liao, W. (2021). Exploring public perceptions of the COVID-19 vaccine online from a cultural perspective: Semantic network analysis of two social media platforms in the United States and China. *Telematics and Informatics, 65*, 101712. <https://doi.org/10.1016/j.tele.2021.101712>
- Lynch, D. C., Greer, A. G., Larson, L. C., Cummings, D. M., Harriett, B. S., Dreyfus, K. S., & Clay, M. C. (2003). Descriptive metaevaluation: Case study of an interdisciplinary curriculum. *Evaluation & the Health Professions, 26*(4), 447–461. <https://doi.org/10.1177/0163278703258099>
- Maboya, M., & McKay, T. (2019). The financial sustainability challenges facing the South African non-profit sector. *The Journal for Transdisciplinary Research in Southern Africa, 15*(1). <https://doi.org/10.4102/td.v15i1.693>
- Mackenzie, M., & Blamey, A. (2005). The practice and the theory: Lessons from the application of a theories of change approach. *Evaluation, 11*(2), 151–168. <https://doi.org/10.1177/1356389005055538>

- Mackewn, J. (2008). Facilitation as action research in the moment. In P. Reason & H. Bradury (Eds.), *The Sage Publications handbook of action research: Participatory inquiry and practice* (2nd ed.) (pp. 615–628). Sage Publications.
- Madill, A., Jordan, A., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology, 91*(1), 1–20. <https://doi.org/10.1348/000712600161646>
- Magnuson, K. A., Sexton, H. R., Davis-Kean, P. E., & Huston, A. C. (2009). Increases in maternal education and young children's language skills. *Merrill-Palmer Quarterly, 55*(3), 319–350. <https://doi.org/10.1353/mpq.0.0024>
- 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), e000617. <https://doi.org/10.1136/bmjgh-2017-000617>
- Markiewicz, A., & Patrick, I. (2016). *Developing monitoring and evaluation frameworks*. Sage Publications.
- Masvaure, S., & Fish, T. E. (2022). Strengthening and measuring monitoring and evaluation capacity in selected African programmes. *African Evaluation Journal, 10*(1), a635. <https://doi.org/10.4102/aej.v10i1.635>
- Masvaure, S., & Motlanthe, S. M. (2022). Reshaping how we think about evaluation: A made in Africa evaluation perspective. *African Evaluation Journal, 10*(1), a618. <https://doi.org/10.4102/aej.v10i1.618>
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Sage Publications.
- Mayne, J. (2007). Challenges and Lessons in implementing results-based management. *Evaluation, 13*(1), 87–109. <https://doi.org/10.1177/1356389007073683>
- Meadows, D. H. (2009). *Thinking in systems: A primer* (D. Wright, Ed.). Earthscan.
- Midgley, G. (2000). *Systemic intervention: Philosophy, methodology, and practice*. Kluwer Academic/Plenum.
- Midgley, G. (2006). Systems thinking for evaluation. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation: An expert anthology* (pp. 11–34). American Evaluation Association.
- Mierlo, B. van. (2010). *Reflexive monitoring in action: A guide for monitoring system innovation projects*. Wageningen UR. <https://research.wur.nl/en/publications/reflexiv-e-monitoring-in-action-a-guide-for-monitoring-system-inno>

- Miller, J. G. (1978). *Living systems*. McGraw-Hill.
- Miller, R. L. (2016). On messes, systems thinking, and evaluation: A response to Patton. *American Journal of Evaluation, 37*(2), 266–269.
<https://doi.org/10.1177/1098214015626294>
- Mingers, J. C. (1980). Towards an appropriate social theory for applied systems thinking: Critical theory and soft systems methodology. *Journal of Applied Systems Analysis, 7*, 41–50.
- Moore, C. G., Carter, R. E., Nietert, P. J., & Stewart, P. W. (2011). Recommendations for planning pilot studies in clinical and translational research. *Clinical and Translational Science, 4*(5), 332–337. <https://doi.org/10.1111/j.1752-8062.2011.00347.x>
- Morgan, D. L. (2007). Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*(1), 48–76. <https://doi.org/10.1177/2345678906292462>
- Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry, 20*(8), 1045–1053. <https://doi.org/10.1177/1077800413513733>
- Moss, E., Patterson, N. A., & Seals, B. F. (2022). An examination of US COVID-19 vaccine distribution in New Jersey, Pennsylvania, and New York. *International Journal of Environmental Research and Public Health, 19*(23), 15629.
<https://doi.org/10.3390/ijerph192315629>
- Mouton, C. (2010). *The history of programme evaluation in South Africa* [Masters Dissertation, University of Stellenbosch]. <https://scholar.sun.ac.za/items/a86ee029-1769-48dc-a055-2a3a9643a275>
- Mowles, C. (2014). Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. *Evaluation, 20*(2), 160–175.
<https://doi.org/10.1177/1356389014527885>
- Mueller-Hirth, N. (2012). If you don't count, you don't count: Monitoring and evaluation in South African NGOs. *Development and Change, 43*(3), 649–670.
<https://doi.org/10.1111/j.1467-7660.2012.01776.x>
- Nayfeld, I., Fuccillo, J., & Greenfield, D. B. (2013). Executive functions in early learning: Extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences, 26*, 81–88.
<https://doi.org/10.1016/j.lindif.2013.04.011>
- Njenga, S. G., & Massyn, L. (2024). Enriching systems practice with African perspectives. *Systems Research and Behavioral Science, 41*(5), 717–725.
<https://doi.org/10.1002/sres.3056>

- OECD. (2019). *Better criteria for better evaluation: Revised evaluation criteria definitions and principles for use*. Organisation of Economic Co-operation and Development. <https://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>
- OECD. (2021). *Applying evaluation criteria thoughtfully*. OECD. <https://doi.org/10.1787/543e84ed-en>
- Ofir, Z. (2017a). Evaluating for development: Reshaping our criteria? *eVALUation Matters, Q1*, 66–79. <https://idev.afdb.org/sites/default/files/Evaluations/2020-02/Evaluating%20for%20Development%20by%20Zenda%20Ofir.pdf>
- Ofir, Z. (2017b, October 6). The DAC criteria, part 2. Why change? [Blog]. *Evaluation for Transformation*. <https://zendaofir.com/updating-the-dac-evaluation-criteria-part-2/>
- Ofir, Z. (2017c, October 16). The DAC criteria, part 5. Non-negotiable criteria [Blog]. *Evaluation for Transformation*. <https://zendaofir.com/updating-dac-evaluation-criteria-part-5/>
- Osman, S., Churruca, K., Ellis, L. A., & Braithwaite, J. (2024). Systems archetypes to investigate the unintended consequences of telehealth in rural Australia: A systems thinking approach to telehealth evaluation and policymaking. *The International Journal of Health Planning and Management, 39*(2), 204–219. <https://doi.org/10.1002/hpm.3741>
- Outcome Mapping Learning Community. (2021). *20 years of Outcome Mapping: Evolving practices for transformative change*. Outcome Mapping Learning Community. <https://www.outcomemapping.ca/resource/20-years-of-outcome-mapping-evolving-practices-for-transformative-change>
- Pan, X., Valerdi, R., & Kang, R. (2013). Systems thinking: A comparison between Chinese and Western approaches. *Procedia Computer Science, 16*, 1027–1035. <https://doi.org/10.1016/j.procs.2013.01.108>
- Parrish, P. E. (2009). Aesthetic principles for instructional design. *Educational Technology Research and Development, 57*(4), 511–528. <https://doi.org/10.1007/s11423-007-9060-7>
- Parsons, B. (2007). *Designing initiative evaluation: A Systems-oriented framework for evaluating social change efforts*. W.K. Kellogg Foundation. <https://www.alnap.org/system/files/content/resource/files/main/designing-initiative-evaluation.pdf>
- Patrizi, P., Heid Thompson, E., Coffman, J., & Beer, T. (2013). Eyes wide open: Learning as strategy under conditions of complexity and uncertainty. *The Foundation Review, 5*(3), 50–65. <https://doi.org/10.9707/1944-5660.1170>

- Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. The Guilford Press.
- Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Sage Publications.
- Patton, M. Q. (Ed.). (2015a). *Developmental evaluation exemplars: Principles in practice*. The Guilford Press.
- Patton, M. Q. (2015b). *Qualitative research & evaluation methods* (4th ed.). Sage Publications.
- Patton, M. Q. (2021). Evaluation criteria for evaluating transformation: Implications for the Coronavirus pandemic and the global climate emergency. *American Journal of Evaluation*, 42(1), 53–89. <https://doi.org/10.1177/1098214020933689>
- Patton, M. Q. (2023). To the editor. *Systems Research and Behavioral Science*, 40(4), 633–635. <https://doi.org/10.1002/sres.2927>
- Patton, M. Q., & Richardson, R. (2024). A philanthropic theory of systems transformation for advancing equity in the polycrisis. *The Foundation Review*, 16(2), 13 – 28. <https://doi.org/10.9707/1944-5660.1706>
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage Publications.
- Peersman, G., & Rogers, P. (2017). *Pathways to advance professionalisation within the context of the Australasian Evaluation Society*. ANZSOG & Better Evaluation. https://www.betterevaluation.org/sites/default/files/Pathways_to_Professionalisation_AES_Final_Public_Version_Peersman_Rogers_reduced.pdf
- Peersman, G., Rogers, P., Guijt, I., Hearn, S., Pasanen, T., & Buffardi, A. L. (2016). *When and how to develop an impact-oriented monitoring and evaluation system*. Overseas Development Institute. <https://odi.org/en/publications/when-and-how-to-develop-an-impact-oriented-monitoring-and-evaluation-system/>
- Peppers, K., Tuunanen, T., Gengler, C. E., Rossi, M., & Hui, W. (2006). The Design Science Research process: A model for producing and presenting Information Systems research. *DESRIST 2006 Proceedings*, 83–106. <https://jyx.jyu.fi/bitstream/handle/123456789/63435/Design%20Science%20Research%20Process.pdf>
- Penna, A. L., de Aquino, C. M., Pinheiro, M. S. N., do Nascimento, R. L. F., Farias-Antúnez, S., Araújo, D. A. B. S., Mita, C., Machado, M. M. T., & Castro, M. C. (2023). Impact of the COVID-19 pandemic on maternal mental health, early childhood development, and parental practices: A global scoping review. *BMC Public Health*, 23, 338. <https://doi.org/10.1186/s12889-023-15003-4>

- Peter, J., Benjamin, P., LeFevre, A. E., Barron, P., & Pillay, Y. (2018). Taking digital health innovation to scale in South Africa: Ten lessons from MomConnect. *BMJ Global Health*, 3(Suppl 2), e000592. <https://doi.org/10.1136/bmjgh-2017-000592>
- Pfeffer, J., & Salancik, G. R. (2009). *The external control of organizations: A resource dependence perspective*. Stanford Business Books.
- Polastro, R. (2014). Evaluating humanitarian action in real time: Recent practices, challenges, and innovations. *Canadian Journal of Program Evaluation*, 29(1), 118-134. <https://doi.org/10.3138/cjpe.29.1.118>
- Preskill, H., & Boyle, S. (2009). Insights into evaluation capacity building: Motivations, strategies, outcomes, and lessons learned. *Canadian Journal of Program Evaluation*, 23(3), 147–174. <https://doi.org/10.3138/cjpe.0023.008>
- Preskill, H., Gopalakrishnan, S., Mack, K., & Cook, J. (2014). *Evaluating complexity: Propositions for improving practice*. Foundation Strategy Group. <https://www.fsg.org/publications/evaluating-complexity#download-area>
- Rajagopalan, R., & Midgley, G. (2015). Knowing differently in systemic intervention. *Systems Research and Behavioral Science*, 32(5), 546–561. <https://doi.org/10.1002/sres.2352>
- Reed, E., Harwood, J., Jackson, J. L. A., & Gebreyesus, T. (2021). “Leveling up” for change: Application of a multilevel system capacity change model to evaluate U.S. Department of Defense global health engagement activities. *New Directions for Evaluation*, 2021(170), 39–50. <https://doi.org/10.1002/ev.20458>
- Reith, S. (2010). Money, power, and donor–NGO partnerships. *Development in Practice*, 20(3), 446–455. <https://doi.org/10.1080/09614521003709932>
- Renger, R. (2015). System evaluation theory (SET): A practical framework for evaluators to meet the challenges of system evaluation. *Evaluation Journal of Australasia*, 15(4), 16–28.
- Renger, R., Atkinson, L., Renger, J., Renger, J., & Hart, G. (2019). The connection between logic models and systems thinking concepts. *Evaluation Journal of Australasia*, 19(2), 79–87. <https://doi.org/10.1177/1035719X19853660>
- Renmans, D., Holvoet, N., & Criel, B. (2020). No mechanism without context: Strengthening the analysis of context in realist evaluations using causal loop diagramming. *New Directions for Evaluation*, 2020(167), 101–114. <https://doi.org/10.1002/ev.20424>
- Republic of South Africa. (2015). *National Integrated Early Childhood Development Policy 2015*. https://www.gov.za/sites/default/files/gcis_document/201610/national-integrated-eed-policy-web-version-final-01-08-2016a.pdf

- Reynolds, M. (2006). Evaluation based on Critical Systems Heuristics. In B. Williams & I. Imam (Eds.), *Systems concepts in evaluation: An expert anthology* (pp. 101–121). American Evaluation Association.
- Reynolds, M., Gates, E., Hummelbrunner, R., Marra, M., & Williams, B. (2016). Towards systemic evaluation. *Systems Research and Behavioral Science*, 33(5), 662–673. <https://doi.org/10.1002/sres.2423>
- Ricciardi, C., Manfra, L., Hartman, S., Bleiker, C., Dineheart, L., & Winsler, A. (2021). School readiness skills at age four predict academic achievement through 5th grade. *Early Childhood Research Quarterly*, 57, 110–120. <https://doi.org/10.1016/j.ecresq.2021.05.006>
- Rice, W. S., Sowman, M. R., & Bavinck, M. (2020). Using theory of change to improve post-2020 conservation: A proposed framework and recommendations for use. *Conservation Science and Practice*, 2(12), e301. <https://doi.org/10.1111/csp2.301>
- Richardson, R. A., & Patton, M. Q. (2021). Leadership-evaluation partnership: Infusing systems principles and complexity concepts for a transformational alliance. *New Directions for Evaluation*, 2021(170), 139–147. <https://doi.org/10.1002/ev.20461>
- Richmond, B. (1994). Systems thinking/system dynamics: Let's just get on with it. *System Dynamics Review*, 10(2–3), 135–157. <https://doi.org/10.1002/sdr.4260100204>
- Ringhofer, L., & Kohlweg, K. (2019). Has the theory of change established itself as the better alternative to the Logical Framework Approach in development cooperation programmes? *Progress in Development Studies*, 19(2), 112–122. <https://doi.org/10.1177/1464993418822882>
- Ríos Romero, M. J., Urquía-Grande, E., & Abril, C. (2023). NGO accountability to donors: Better said than done? *Revista de Estudios Empresariales. Segunda Época*, 31–51. <https://doi.org/10.17561/ree.n1.2023.7410>
- Ritchey, T. (2013). Wicked problems: Modelling social messes with morphological analysis. *Acta Morphologica Generalis*, 2(1), 9.
- Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155–169. <https://doi.org/10.1007/BF01405730>
- Rogers, M. (2015). Contextualizing theories and practices of bricolage research. *The Qualitative Report*, 17, 1–17. <https://doi.org/10.46743/2160-3715/2012.1704>
- Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1), 29–48. <https://doi.org/10.1177/1356389007084674>
- Rogers, P., & Macfarlan, A. (2020). *An overview of monitoring and evaluation for adaptive management*. BetterEvaluation.

https://www.betterevaluation.org/sites/default/files/MandE_for_adaptive_management_WP1_Overview_202009.pdf

- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). *Evaluation: A systematic approach* (8th ed.). Sage Publications.
- Salter, K. L., & Kothari, A. (2014). Using realist evaluation to open the black box of knowledge translation: A state-of-the-art review. *Implementation Science*, *9*(1), 115. <https://doi.org/10.1186/s13012-014-0115-y>
- Schmidt-Abbey, B. (2023, March 29). "In Search of a Golden Mean for Systemic Evaluation"—A public debate between Michael Quinn Patton and Mike C. Jackson hosted by EES-EES [Blog]. *TWG8 Blog*. <https://europeanevaluation.org/2023/03/29/in-search-of-a-golden-mean-for-systemic-evaluation-a-public-debate-between-michael-quinn-patton-and-mike-c-jackson-hosted-by-ees/>
- Seel, N. M., Lehmann, T., Blumschein, P., & Podolskiy, O. A. (2017). *Instructional design for learning: Theoretical foundations*. Sense Publishers.
- Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization* (Revised). Currency.
- Senge, P. M. (2006). *The fifth discipline: The art and practice of the learning organization* (2nd ed.). Doubleday.
- SETIG. (2018). *Principles for effective use of systems thinking in evaluation practice*. Systems in Evaluation Topical Interest Group of the American Evaluation Association. <https://www.betterevaluation.org/sites/default/files/SETIG-Principles-FINAL-DRAFT-2018-9-9.pdf>
- Sgourou, E., Katsakiori, P., Papaioannou, I., Goutsos, S., & Adamides, E. (2012). Using soft systems methodology as a systemic approach to safety performance evaluation. *Procedia Engineering*, *45*, 185–193. <https://doi.org/10.1016/j.proeng.2012.08.141>
- Shearn, K., Allmark, P., Piercy, H., & Hirst, J. (2017). Building realist program theory for large complex and messy interventions. *International Journal of Qualitative Methods*, *16*(1), 160940691774179. <https://doi.org/10.1177/1609406917741796>
- Shiell, A., Hawe, P., & Gold, L. (2008). Complex interventions or complex systems? Implications for health economic evaluation. *BMJ*, *336*(7656), 1281–1283. <https://doi.org/10.1136/bmj.39569.510521.AD>
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697. <https://doi.org/10.1177/0956797612440101>

- Simister, N. (2017). *Results-based management*. INTRAC. <https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Results-based-Management.pdf>
- Smith, R., Mauremootoo, J., & Rassmann, K. (2012). *Ten years of Outcome Mapping adaptations and support*. Outcome Mapping Learning Community. https://www.outcomemapping.ca/download/10_years_of_OM.pdf
- Snelling, M., Dawes, A., Biersteker, L., Girdwood, E., & Tredoux, C. (2019a). *Early Learning Outcomes Measure technical manual* [dataset]. Innovation Edge. <https://doi.org/10.1037/t71873-000>
- Snelling, M., Dawes, A., Biersteker, L., Girdwood, E., & Tredoux, C. (2019b). The development of a South African early learning outcomes measure: A South African instrument for measuring early learning program outcomes. *Child: Care, Health and Development*, 45(2), 257–270. <https://doi.org/10.1111/cch.12641>
- Snowden, D. (2005). Strategy in the context of uncertainty. *Handbook of Business Strategy*, 6(1), 47 – 54. <https://doi.org/10.1108/08944310510556955>
- Snowden, D. J., & Boone, M. E. (2007). A leader's framework for decision making. *Harvard business review*, 85(11), 68 – 76.
- South African Government. (2018, May 30). *Western Cape Social Development hosts Early Childhood Development organisations as part of Child Protection Week*. <https://www.gov.za/news/media-statements/western-cape-social-development-hosts-early-childhood-development>
- Spaull, N., & Van Der Berg, S. (2020). Counting the cost: COVID-19 school closures in South Africa and its impact on children. *South African Journal of Childhood Education*, 10(1). <https://doi.org/10.4102/sajce.v10i1.924>
- Stacey, R. D. (1992). *Managing the unknowable: Strategic boundaries between order and chaos in organizations*. Jossey-Bass.
- Stephens, A., Lewis, E.D. and Reddy, S.M. (2018). *Inclusive Systemic Evaluation (ISE4GEMs): A new approach for the SDG era*. UN Women.
- Sterman, J. D. (2001). System dynamics modeling: Tools for learning in a complex world. *California Management Review*, 4, 8–25.
- Stockdill, S. H., Baizerman, M., & Compton, D. W. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, 2002(93), 7–26. <https://doi.org/10.1002/ev.39>
- Stroh, D. P. (2015). *Systems thinking for social change*. Chelsea Green Publishing.

- Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation*, 2001(89), 7–98. <https://doi.org/10.1002/ev.3>
- Stufflebeam, D. L., & Coryn, C. L. S. (Eds.). (2014). *Evaluation theory, models, & applications* (2nd ed.). Jossey-Bass.
- Swanson, D. J., & Creed, A. S. (2014). Sharpening the Focus of Force Field Analysis. *Journal of Change Management*, 14(1), 28–47. <https://doi.org/10.1080/14697017.2013.788052>
- Taylor-Powell, E., & Boyd, H. H. (2008). Evaluation capacity building in complex organizations. *New Directions for Evaluation*, 2008(120), 55–69. <https://doi.org/10.1002/ev.276>
- Tebes, J. K. (2005). Community science, philosophy of science, and the practice of research. *American Journal of Community Psychology*, 35(3–4), 213–230. <https://doi.org/10.1007/s10464-005-3399-x>
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Sage Publications.
- Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. In C. Willig & W. Stainton-Rogers (Eds.), *The Sage Publications Handbook of Qualitative Research in Psychology* (2nd ed., pp. 17–37). Sage Publications.
- Tredoux, C., Dawes, A., Mattes, F., Schenk, J.-C., Giese, S., Leach, G., Van Der Berg, S., & Horler, J. (2024). Are South African children on track for early learning? Findings from the South African Thrive By Five Index 2021 Survey. *Child Indicators Research*, 17(2), 601–636. <https://doi.org/10.1007/s12187-023-10093-3>
- Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation*, 123, 13 – 32. <https://doi.org/10.1002/ev.303>
- Trochim, W., Urban, J. B., Hargreaves, M., Hebbard, C., Buckley, J., Archibald, T., Johnson, M., & Burgermaster, M. (2016). *The guide to the Systems Evaluation Protocol (V3.1)*. Cornell Office for Research on Evaluation (CORE), Cornell University, and Research on Evaluation and Developmental Systems Science Lab (REDSS Lab), Montclair State University. <https://core.human.cornell.edu/research/systems/protocol/index.cfm>
- Uitto, J. I. (2021). Confronting storms, fires and pestilence: Meaningful evaluation for a hazardous world. *Evaluation Matters—He Take Tō Te Aromatawai*, 7, 96–126. <https://doi.org/10.18296/em.0068>
- Ulrich, W. (1983). *Critical heuristics of social planning: A new approach to practical philosophy*. Haupt.
- Ulrich, W. (1987). Critical heuristics of social systems design. *European Journal of Operational Research*, 31(3), 276–283. [https://doi.org/10.1016/0377-2217\(87\)90036-1](https://doi.org/10.1016/0377-2217(87)90036-1)

- Ulrich, W. (1988). Churchman's 'process of unfolding'- Its significance for policy analysis and evaluation. *Systems Practice*, 1(4), 415–428. <https://doi.org/10.1007/BF01066583>
- Ulrich, W., & Reynolds, M. (2010). Critical systems heuristics. In M. Reynolds & S. Holwell (Eds.), *Systems approaches to managing change: A practical guide* (pp. 243–292). Springer.
- UNDG. (2011). *Results-Based management handbook: Harmonizing RBM concepts and approaches for improved development results at country level*. United Nations Development Group. <https://unsdg.un.org/sites/default/files/UNDG-RBM-Handbook-2012.pdf>
- United Way of America. (1996). *Measuring program outcomes: A practical approach*. United Way of America.
- Urban, J. B., Hargraves, M., Buckley, J., Archibald, T., Hebbard, C., & Trochim, W. M. (2021). The Systems Evaluation Protocol for evaluation planning. *New Directions for Evaluation*, 2021(169), 31–45. <https://doi.org/10.1002/ev.20443>
- Urwin, E., Botoeva, A., Arias, R., Vargas, O., & Firchow, P. (2023). Flipping the power dynamics in measurement and evaluation: International aid and the potential for a grounded accountability model. *Negotiation Journal*, 39(4), 401–426. <https://doi.org/10.1111/nejo.12448>
- USAID. (2020). *A guide to complexity-aware monitoring approaches for MOMENTUM projects*. USAID. <https://usaidmomentum.org/resource/a-guide-to-complexity-aware-monitoring-approaches-for-momentum-projects/>
- Valters, C. (2014). *Theories of Change in international development: Communication, learning, or accountability?* The Asia Foundation. <https://usaidlearninglab.org/system/files/resource/files/theoriesofchange.pdf>
- Vereijken, B. (2010). The complexity of childhood development: Variability in perspective. *Physical Therapy*, 90(12), 1850–1859. <https://doi.org/10.2522/ptj.20100019>
- Visser, M., Juan, A., & Feza, N. (2015). Home and school resources as predictors of mathematics performance in South Africa. *South African Journal of Education*, 35(1), 1–10. <https://doi.org/10.15700/201503062354>
- Vogel, I. (2012). *Review of the use of 'Theory of Change' in international development*. UK Department for International Development (DFID). https://assets.publishing.service.gov.uk/media/57a08a5ded915d3cfd00071a/DFID_ToC_Review_VogelV7.pdf
- Von Bertalanffy, L. (1950). The theory of open systems in physics and biology. *Science*, 111(2872), 23–29. <https://doi.org/10.1126/science.111.2872.23>

- Vorwerk Marren, I., Davis, A., & Williamson, C. M. (2024). Strategizing for survival – enablers of South African not-for-profit organization sustainability. *Cogent Business & Management*, 11(1), 2323775. <https://doi.org/10.1080/23311975.2024.2323775>
- W. K. Kellogg Foundation. (2004). *Logic model development guide*. W. K. Kellogg Foundation. <https://www.wkkf.org/resource-directory/resources/2004/01/logic-model-development-guide>
- Wade, J., & Kallemeyn, L. (2020). Evaluation capacity building (ECB) interventions and the development of sustainable evaluation practice: An exploratory study. *Evaluation and Program Planning*, 79, 101777. <https://doi.org/10.1016/j.evalprogplan.2019.101777>
- Walker, B., Gunderson, L., Kinzig, A., Folke, C., Carpenter, S., & Schultz, L. (2006). A handful of heuristics and some propositions for understanding resilience in social-ecological systems. *Ecology and Society*, 11(1), art13. <https://doi.org/10.5751/ES-01530-110113>
- Walton, M. (2016). Expert views on applying complexity theory in evaluation: Opportunities and barriers. *Evaluation*, 22(4), 410–423. <https://doi.org/10.1177/1356389016667890>
- Walton, M., & Baker, V. (2021). The role of evaluation in supporting action on a hazardous world, a call for critical systems thinking. *Evaluation Matters—He Take Tō Te Aromatawai*, 7, 148–157. <https://doi.org/10.18296/em.0071>
- Walton, M., Gates, E. F., & Vidueira, P. (2021). Insights and future directions for systems and complexity-informed evaluation. *New Directions for Evaluation*, 2021(170), 159–171. <https://doi.org/10.1002/ev.20459>
- Wasserman, D. L. (2010). Using a systems orientation and foundational theory to enhance theory-driven human service program evaluations. *Evaluation and Program Planning*, 33(2), 67–80. <https://doi.org/10.1016/j.evalprogplan.2009.06.005>
- Weaver, W. (1948). Science and complexity. *American Scientist*, 36, 536–544.
- Weiner, N. (1948). *Cybernetics*. MIT Press.
- Weiss, C. H. (1997). Theory-based evaluation: Past, present, and future. *New Directions for Evaluation*, 1997(76), 41–55. <https://doi.org/10.1002/ev.1086>
- Westhorp, G. (2014). *Realist impact evaluation: An introduction*. Overseas Development Institute. <https://cdn.odi.org/media/documents/9138.pdf>
- Westley, F., Patton, M. Q., & Zimmerman, B. (2011). *Getting to maybe: How the world is changed*. Vintage Canada.

- Whitty, B. (2013). *Experiences of the results agenda*. Big Push Forward.
<https://bigpushforward.com/wp-content/uploads/2011/01/Experiences-of-the-Results-Agenda-by-Brendan-Whitty.pdf>
- Wholey, J. S. (1979). *Evaluation: Promise and performance*. Urban Institute.
- Wholey, J. S. (2001). Managing for results: Roles for evaluators in a new management era. *American Journal of Evaluation*, 22(3), 343–347.
- Wildschut, L. P. (2014). *Theory-based evaluation, logic modelling and the experience of South African non-governmental organisations* [Unpublished doctoral dissertation]. Stellenbosch University.
- Wilkinson, H., Hills, D., Penn, A., & Barbrook-Johnson, P. (2021). Building a system-based Theory of Change using Participatory Systems Mapping. *Evaluation*, 27(1), 80–101.
<https://doi.org/10.1177/1356389020980493>
- Williams, B. (2019). *Systemic evaluation design: A workbook* (2nd ed.). Gumroad.
- Williams, B., & Hummelbrunner, R. (2011). *Systems concepts in action: A practitioner's toolkit*. Stanford University Press.
- Williams, B., & Imam, I. (Eds.). (2006). *Systems concepts in evaluation: An expert anthology*. American Evaluation Association.
- Willig, C. (2013). *Introducing qualitative research in psychology* (3rd ed). McGraw-Hill Open University Press.
- Wills, G., & Kika-Mistry, J. (2021). *Early Childhood Development in South Africa during the COVID-19 pandemic: Evidence from NIDS-CRAM waves 2—5*. National Income Dynamics Study. https://cramsurvey.org/wp-content/uploads/2021/07/14.-Wills-G_-Kika-Mistry-J.-2021-Early-Childhood-Development-in-South-Africa-during-the-n-COVID-19-pandemic- Evidence-from-NIDS-CRAM-Waves-2-5.pdf
- Witesman, E., & Heiss, A. (2017). Nonprofit collaboration and the resurrection of market failure: How a resource-sharing environment can suppress social objectives. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 28(4), 1500–1528. <https://doi.org/10.1007/s11266-016-9684-5>
- Wolf-Branigin, M. (2013). *Using complexity theory for research and program evaluation*. Oxford University Press.
- Yeboah, S. A. (2022). Solving local problems or looking good: An ethnography of the field practices of foreign sponsored ngos in rural african communities. *The European Journal of Development Research*, 34(3), 1645–1661. <https://doi.org/10.1057/s41287-021-00442-9>

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.

Zazueta, A. E., Le, T. T., & Bahramalian, N. (2021). Development trajectories and complex systems–informed theories of change. *American Journal of Evaluation*, 42(1), 110–129. <https://doi.org/10.1177/1098214020947782>

Zimmerman, B., Lindberg, C., & Plsek, P. (1998). *Edgework: Lessons from complexity science for health care leaders*. VHA Incorporated.

Appendix A: Letters of Invitation to Expert Reviewers

Note that some personal information has been removed to protect the identity of the reviewers and pseudonyms are used.

International Expert

Dear Sarah

I am currently doing my PhD in programme evaluation at the University of Cape Town in South Africa, under the supervision of [Assoc Prof Sarah Chapman](#) and [Dr Carren Duffy](#).

My dissertation involves the creation and pilot of a MEL systems thinking toolkit that local evaluators can implement with their nonprofit clients to create a systemic MEL framework.

I work in MEL consulting, primarily in the Early Childhood Development (ECD) sector, and am motivated to upskill local practitioners and ECD nonprofits to think more systemically about their programmes. However, practical and intellectual access to the systems thinking literature is difficult to navigate in our context, particularly within nonprofit teams (many of whom are not tertiary educated). Hence the drive to create this accessible and user-friendly toolkit that provides a way to benefit from the field without needing a degree in cybernetics!

My understanding of this topic has been heavily influenced by your work, particularly your earlier papers from [...], and of course the [...].

For this reason, I am writing to gauge your interest and willingness to participate in an expert review of the toolkit. My fieldwork involves an initial expert review, followed by a pilot in which it is implemented in local ECD nonprofits. After each step, I will incorporate the expert and user feedback into the various steps and tools for overall improvement. Given your expertise within this field, as well as your background in educational psychology, you are the first international reviewer that came to mind!

The review involves reading through the toolkit - which is 189 pages long (a large chunk of this is references, worksheets and templates) - and then providing structured written feedback by the end of August. I estimate this to take a total of about two days/16 hours.

I have no doubt that you are incredibly busy with your teaching and research, and I am asking for a significant commitment, however your insights and feedback would be a massive contribution to this academic work and getting systems thinking into the hands of South African evaluators.

If this is something that you are interested in doing, I can send you more information or we can set up a call.

I look forward to hearing from you.

Warm regards,

Jessica Horler

Local Experts

Dear Alletta / Naledi

I am currently pursuing my PhD in programme evaluation at UCT, under the supervision of A/Prof Sarah Chapman and Dr Carren Duffy. My dissertation involves the creation and pilot of a MEL systems thinking toolkit that local evaluators can implement with their nonprofit clients to create a systemic MEL framework. I have been doing MEL consulting for a number of years, and more recently, been focusing on the Early Childhood Development sector. As such, I am motivated to use this PhD work to upskill local evaluation practitioners and ECD nonprofits to think more systemically about their programmes, using this an accessible and user-friendly MEL toolkit.

I have developed a draft toolkit to be used by MEL practitioners in the ECD sector, and the bulk of my PhD research will involve documenting the applicability of this toolkit to MEL practitioners, and also then refining and adapting the toolkit based on this feedback. One of the steps worked into the methodology at this stage is an “expert review” of the draft toolkit, to give me feedback on its coherence and feasibility. We were wondering therefore on your willingness to participate in a review of the toolkit? Your evaluation experience in the education sector would be incredibly valuable to this process, as well as your experience facilitating the development of MEL frameworks.

The review involves reading through the toolkit - which is 189 pages long (a large chunk of this is references, worksheets and templates) - and then providing structured written feedback by the end of August. We would be able to offer an honorarium of R5000 for the review work.

I have no doubt that you are incredibly busy – but please let me know if this is something you would be interested in contributing toward.

Warm regards,

Jessica

Appendix B: Reviewer Consent Form

Consent Form: Expert Review

Study: A systems-based approach to strategic planning, monitoring and evaluation: Designing and testing the SAMEL Toolkit

Researcher: Jessica Horler

jesshorler@gmail.com ; +27824675760



Dear

You have been invited to participate in a doctoral research project conducted by Jessica Horler from the University of Cape Town. This project will design and test a systems thinking toolkit for the development of a Monitoring, Evaluation and Learning (MEL) framework for use by non-profit organisations in the Early Childhood Development sector (the 'SAMEL Toolkit').

You have opted to take part in the expert review of this toolkit. The objective of this process is to sense-check the content of the toolkit in line with your subject matter expertise in MEL. Your participation in this review involves reading through all toolkit documentation and providing written feedback via the document provided. Your feedback may be referenced in the dissertation. The review should take you an estimated two days, and you will receive an honorarium of R5000.

Please take note of the following:

- Your participation is voluntary.
- You have the right to withdraw from the review at any time.
- Information that you provide as part of this review will be subject to standard data use policies which protect anonymity of individuals and institutions. The feedback that you provide will be reported on anonymously.
- Your expertise and experience will be described (anonymously) in the dissertation to explain your selection as a reviewer.
- All contents of the toolkit are the intellectual property of the University of Cape Town. By signing below, you agree not to share, copy, publish, replicate or reference any of the materials in your personal or professional capacity until the dissertation is completed and published.

Please indicate your consent for this interview by signing below.

I, the undersigned, have read and acknowledged the content conditions above, and consent to take part in the review.

Name

Signature

Date

DD

M

YY

M

Appendix C: Introductory Letter to Toolkit (Reviewers)



Dear

Once again, thank you for agreeing to take part in this review.

Some additional background on this project:

I have been consulting in MEL since completing my master's degree in 2015 (in Programme Evaluation). In the last four years, I have begun specialising in Early Childhood Development (ECD) programmes. My consulting work typically entails programme theory development, MEL system frameworks, process and outcome evaluations, and other desktop research.

My original intention with my PhD was to look at how NGO leadership and organisational dynamics impact programme effectiveness. I registered for the degree in February 2020, a month before the world went into lockdown. After working on this topic for a year, I had to let it go as it necessitated fieldwork that just wasn't possible in the foreseeable future. However, I was still intrigued by the contextual messiness of social intervention, and how this comes to bear on MEL. I dove deeply into the systems thinking and complexity science literature, and of course, came across your work which was hugely beneficial with regards to understanding the intersection between these fields and evaluation.

It struck me how useful and applicable these fields are, not just to evaluation and evaluators, but to programme monitoring and learning for programme managers. However, I could not find any comprehensive guidelines, books or documents that would help your average, non-academic evaluator apply these concepts to their work with nonprofits, specifically around programme strategy development and learning. Hence this toolkit.

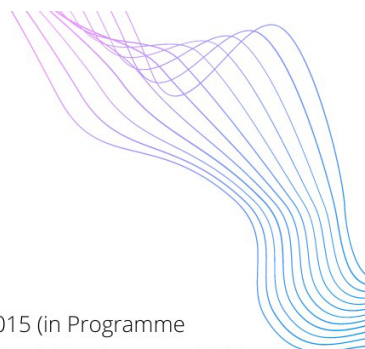
The toolkit has evolved over the last year to be as streamlined and simplified as possible so as to appeal to non-academic evaluators (it was originally much longer and much heavier on the systems thinking side), particularly in South Africa where a substantial chunk of the evaluator pool does not have a master's degree.

Following the review, the toolkit is going to be piloted at two local ECD organisations. The goal of this review is to have it sense-checked by experts, and to make changes based on this expert feedback before going into the field. Thereafter, its practical utility, relevance and feasibility will be assessed in the pilot phase.

The review questions are contained in the Google Drive folder. Please provide your written feedback in the document, but also please feel free to leave comments in the PDF chapters if you want to.

I am incredibly grateful for your contribution, and really look forward to your feedback.

*Thank you!
Jess*



Appendix D: Expert Review Form

The review questions are contained in the table on the following page. Please provide your feedback in the labelled column.

These questions are based on the general goals of the SAMEL Toolkit (listed following the review guidelines). These goals are derived from the African Evaluation Guidelines 2020¹², and the Made in Africa Evaluation Principles by Chilisa and Mertens (2021)¹³. The goals of the toolkit were aligned with these documents in order to (1) contribute to a lacking pool of African-made evaluation models, frameworks and resources, and (2) ensure that the toolkit acknowledges and advances its African context.

¹² <https://afrea.org/wp-content/uploads/2020/03/AEG-29-February-2020-FINAL-DRAFT-for-consultation.pdf>

¹³ Chilisa, B., & Mertens, D. M. (2021). Indigenous Made in Africa Evaluation frameworks: Addressing epistemic violence and contributing to social transformation. *American Journal of Evaluation*, 42(2), 241–253. <https://doi.org/10.1177/1098214020948601>

Goals	Review Questions	Feedback
<p><i>The toolkit empowers the evaluator and the NPO team.</i></p>	<p>1. Please comment on this assertion: “the toolkit facilitates discovery, learning and agency for <i>both</i> the evaluator and the team”</p>	
	<p>2. To what extent do you think the toolkit can add to an evaluator’s MEL competencies? Is there any specific component of the toolkit that you think is particularly valuable?</p>	
	<p>3. Is the toolkit likely to develop evaluative and learning capacity within an NPO team? Please elaborate.</p>	
	<p>4. Do you think the workshop guidelines (included at the end of each chapter) are sufficient for stimulating the programme team’s participation and input? If not, how can this be improved?</p>	
	<p>5. Are the toolkit steps likely to prepare the team to plan, adjust, be accountable and use programme data? Are certain steps or tools more effective than others in this vein?</p>	
	<p>6. How feasible are the toolkit steps (including preparation and workshop facilitations) given typical evaluator competencies (in South Africa/the developing world)? Are some steps or tools less accessible/user-friendly than others?</p>	
	<p>7. How appropriate/user-friendly do you think the toolkit steps are to an NPO team? Are some steps less appropriate than others?</p>	
<p><i>The toolkit is technically robust.</i></p>	<p>8. Do you think that the toolkit is sufficient to support evaluators and NPO teams to co-create a technically rigorous MEL framework? Why or why not?</p>	
	<p>9. Were any of the steps missing important sources, evidence, concepts or tools?</p>	
	<p>10. How feasible are the toolkit steps in terms of typical NPO resources? (human, financial, capacity and time)</p>	
<p><i>The toolkit is ethically sound.</i></p>	<p>11. Do the toolkit steps and tools enable multiple perspectives to be incorporated into the outputs,</p>	

	including those of marginalised groups outside of the programme?	
<i>The toolkit has value for the evaluation field.</i>	12. Do you think that there is a need for this toolkit? Please explain.	
	13. What value do you think this toolkit has for both the evaluation and ECD communities? (e.g., contribution to programme effectiveness, increasing evaluative capacity, mainstreaming systems thinking in evaluation)	
	14. Are certain sections or outputs of the toolkit more valuable than others? Please explain.	
	15. Do you think that this toolkit sufficiently reflects a systems thinking practice (keeping in mind its purpose and target audience)?	
<i>Other comments</i>	16. Do you have any other comments on the toolkit?	

TOOLKIT GOALS

1. The toolkit empowers the evaluator and the NPO team.
 - a. It facilitates discovery, learning and agency among the evaluator and the team.
 - b. It develops capacity in evaluative thinking among the evaluator and the team.
 - c. It values the team's contributions.
 - d. It provides the team with the insights to plan, adjust, be accountable and use the programme results.
 - e. It is accessible to a wide evaluator and NPO audience.
2. The toolkit is technically robust.
 - a. It is relevant for the needs and priorities of South African evaluators.
 - b. It is relevant for the needs and priorities of South African ECD NPOs.
 - c. It is coherent.
 - d. It empowers the evaluator and team to co-create a technically rigorous MEL framework.
 - e. It draws on defensible evidence from diverse sources.
 - f. It supports the development of a MEL framework that is based on defensible evidence from diverse sources.
 - g. It encourages the surfacing of team members' own pre-dispositions (values, worldviews, assumptions, etc.)
 - h. It encourages the surfacing of the evaluator's own pre-dispositions.
 - i. It is feasible to implement the toolkit in light of typical evaluator competencies.
 - j. It is feasible to implement the toolkit in light of typical NPO competencies and resources.
 - k. It reflects the context and culture of South African ECD NPOs.
3. The toolkit is ethically sound.
 - a. It respects the knowledge and resources provided by the team members.
 - b. It is inclusive of relevant perspectives, including those of marginalised groups outside of the programme team.
 - c. It minimises power imbalances between the evaluator and team.
4. The toolkit is rooted in Africa, but draws from across the world.
 - a. The products of the toolkit can be used to contribute to transformation and healing within the ECD sector.
 - b. It brings together international and local knowledge.
5. The toolkit reflects the connectedness of the world.
 - a. It supports evaluators to understand the importance of interdependence, relationships and collectiveness, and to incorporate these factors into their practice.
 - b. The toolkit products reflect the relationships important to the programme.
 - c. Its products support sustainable positive change within the NPO, and the NPO's ability to create enduring change within the ECD sector.

**INVITATION TO PARTICIPATE IN DOCTORAL RESEARCH
UNIVERSITY OF CAPE TOWN**

**A SYSTEMS-BASED
APPROACH TO STRATEGIC
MONITORING,
EVALUATION & LEARNING**



This research study is pilot testing a participatory approach to the development of a monitoring, evaluation and learning (MEL) framework for local early childhood development (ECD) programmes.

WHAT IS THIS RESEARCH ABOUT?

There is an ever-growing need to grapple with the complex social conditions in which non-profits intervene, and to create lasting, systemic impacts. As such, we need innovative tools to strategically assist and meet these goals.

We have used systems thinking concepts to help design a new approach to guide monitoring, evaluation and learning in South African ECD programmes. Our approach will involve an experienced evaluator working with your organisation through a series of activities, worksheets and templates completed over a four-day workshop. The output will be a customized monitoring, evaluation and learning framework for your organisation.

We are looking for two ECD organisations to take part: one in an initial pilot (which will be offered to you for free); and a second phase (offered at a discounted rate).

WHO IS ELIGIBLE TO PARTICIPATE?

To participate in this study, your organisation should be based in the Western Cape. Your programme should be situated within the ECD sector, broadly aimed at the improvement of young children's situations. You should have an established programme team and the capacity (staff and time) to develop and implement an MEL framework. You do not have to be well-versed in MEL or have prior experience in MEL.

CONTACT INFORMATION:

Jessica Horler
0824675760
systemsevalstudy@gmail.com

*This study has been approved
by the UCT Commerce
Faculty's Ethics in Research
Committee.*

INVITATION TO PARTICIPATE IN DOCTORAL RESEARCH UNIVERSITY OF CAPE TOWN

HOW WOULD WE BENEFIT?

During the course of the workshops, your programme team will: learn to approach your programme strategy in a systemic way; learn more about systems thinking and related tools and concepts; and design an MEL framework that is responsive to the complex context in which your programme operates.

The framework will consist of the following outputs, designed around your programme and co-created with your team: a situation analysis; programme scope, Theory of Change, Theory of Action, monitoring plan, learning strategy, and evaluation plan.

WHAT WOULD BE REQUIRED OF US?

You and your team need to be available to partake in the study between October and November 2022 for the pilot, or January and March for the second (payable) phase.

Participation will include four full-day workshops. Depending on your time requirements, the workshops can be split into half-day sessions as needed. Your team will also participate in post-workshop focus groups (after each one), and a follow-up interview with the programme director a month after the workshops are completed. All of these activities will be scheduled at your convenience. Workshops will be observed and documented by a fieldworker, and recorded.

WHAT ELSE DO WE NEED TO KNOW?

This research involves field testing this toolkit in terms of its feasibility, usability and effectiveness. While the toolkit has been reviewed by expert evaluators, it has not been tested yet with an organisation. As such, teams should be mindful of this and be willing to participate in a larger learning process for both your organisation, the researcher, and the evaluation community as a whole.

If you are interested in participating in this research please complete [this online form](#).

CONTACT INFORMATION:

Jessica Horler

0824675760

systemsevalstudy@gmail.com

*This study has been approved
by the UCT Commerce
Faculty's Ethics in Research
Committee.*

Appendix F: NPO Eligibility Form

Participation in doctoral research: Eligibility form

Thank you for your interest in participating in the doctoral study: A systems-based approach to strategic monitoring, evaluation and learning.

Participating in this research requires developing a monitoring, evaluation and learning (MEL) framework for an ECD-related programme. This form assesses your eligibility for the study by asking questions about your programme.

Please note that the information you provide will be kept in the strictest confidence, stored securely and not shared with any third party. This information will not be used for any purposes beyond selecting eligible cases for this research study.

** Indicates required question*

1. Name of your organisation *

2. Please provide a link to your website *

3. Your name and role in your organisation (e.g. executive director, programme manager) *

4. Your email address and phone number *

5. Name of your ECD programme *

6. Please select your programme's area/s of focus. (if other, please specify) *

Tick all that apply.

- Nutritional support
- Maternal health
- Child health
- Social services
- Support for primary caregivers
- Early learning
- Resources and training
- Advocacy
- Other: _____

7. Please select your programme's target population (those who you work with DIRECTLY) (if other, please specify) *

Tick all that apply.

- Children
- Caregivers
- ECD practitioners
- Communities
- Sector stakeholders (e.g. government)
- Other: _____

8. What year was your organisation established? *

9. How long has your programme been running? *

Tick all that apply.

- Less than 1 year
- 1 to 2 years
- 3 to 5 years
- More than 5 years

10. Do you and your team have any experience with monitoring and evaluation? *

Mark only one oval.

- A lot
- Some
- Very little
- Not at all

11. If yes, please select if your programme has any of the following:

Tick all that apply.

- Situation analysis or problem statement
- Theory of Change
- Theory of Action
- Logical framework
- Indicators
- Outcome measures (e.g. questionnaires)
- Other: _____

12. Does your organisation have an in-house MEL officer or specialist? *

Mark only one oval.

- Yes
- No

13. Do you have a budget for MEL? *

Mark only one oval.

- Yes
- No

14. Has your programme ever undergone a summative evaluation? (Investigating participant outcomes or impact) *

Mark only one oval.

- Yes
- No

15. Please describe the structure of your programme team - i.e. the number of staff and their roles *

16. Do you have the resources and capacity to implement an MEL framework? (staff time, expertise and funding) *

Mark only one oval.

- Yes
- No
- Not sure - I don't know what resources and capacity are needed.

17. Do you think your organisation has a learning culture? *

Mark only one oval.

- Yes
- No
- I don't know what that is

18. If yes, how would you describe your learning culture?

19. Can your programme team commit to the following between October and November 2022, or January and March 2023: completion of an online survey by all team members; participation in a four-day workshop, participation in a short focus group following each workshop, a follow-up interview with the programme manager or executive director in the month after the workshops. *

Tick all that apply.

Yes

No

20. Please select your preference for the free pilot (Oct/Nov 22) or the second, payable phase (Jan 23). *


Tick all that apply.

2022

2023

Thank you for taking the time to complete this form! If you are eligible to participate in this research you will be contacted in the next few weeks.


Appendix G: Letter to Participate in Pilot



Consent Form: Programme Director

Study: A systems-based approach to strategic planning, monitoring and evaluation in ECD: Designing and testing the SAMEL Toolkit

Researcher: Jessica Horler



You have been invited to participate in a doctoral research project conducted by Jessica Horler from the University of Cape Town. This project will design and test a systems-based toolkit for the development of a monitoring, evaluation and learning (MEL) framework for use by nonprofit organisations in the Early Childhood Development sector.

Your organisation has been selected to take part in the pilot testing of this protocol. The objective of this process is to run through the toolkit via five workshops. During these workshops, you and your team will collaborate with the researcher to develop a MEL framework for your programme.

Your participation in this pilot involves: providing programme documents to the researcher in advance of the workshops; collaborating with the researcher to arrange a date for the workshops; arranging your team's participation; and actively engaging in all of the workshops and workshop preparation.

It also involves providing in-person feedback regarding your experience of the workshop, following each one and two months after the workshops have concluded.

This feedback will be obtained from you in focus groups following each workshop; and in an one-on-one interview two months later.

Please take note of the following:

- Your participation in this pilot is voluntary.
- You have the right to withdraw from the pilot at any time.
- The workshops, focus groups and interview will be audio-recorded and transcribed.
- You, and your organisation, will **not** be identified by name in the dissertation report, nor any associated reports.
- Information that you provide as part of this research will be subject to standard data use policies which protect anonymity of individuals and institutions.
- The workshop and any associated reports provided to you will be free-of-charge.
- The decision to include certain members of your programme team in the workshop is yours to make. The researcher will provide you with guidance to make these decisions.

Please indicate your consent for this research by responding to the conditions below:

<input type="checkbox"/> I understand the purpose of the research	<input type="checkbox"/> I understand that my feedback will be recorded so that the researcher may accurately portray my responses
<input type="checkbox"/> I understand that I can withdraw from this research at any time without giving any reason	<input type="checkbox"/> I understand that my identity (and that of my organisation) will remain confidential and, where appropriate, a pseudonym will be used in referring to information that I provide
<input type="checkbox"/> I understand that the information that I provide may be included in an institutional report	<input type="checkbox"/> I consent to take part in this research.

Name	<input style="width: 150px; height: 20px;" type="text"/>	<input style="width: 150px; height: 20px;" type="text"/>	<input style="width: 150px; height: 20px;" type="text"/>
	First Name	Middle Name	Last Name
Signature	<input style="width: 200px; height: 20px;" type="text"/>		Date
			<input style="width: 60px; height: 20px;" type="text"/> <input style="width: 60px; height: 20px;" type="text"/> <input style="width: 60px; height: 20px;" type="text"/>
			MM DD YY

Appendix H: Focus Group Schedules

FG 1: Scoping Workshop

Set-up and Introduction	
<ul style="list-style-type: none"> Begin by thanking the team for participating in the first SAMEL workshop and for taking the time to provide feedback in this session. Introduce the purpose of the focus group – to obtain feedback and suggestions for improving the workshop protocol. Stress that there are no right or wrong answers, and they need not hold back on constructive criticism. Explain that these focus groups will take place after each of the workshops. 	
Focus Group Questions	
Questions	Probes
1. It would be great to begin all of these reflection sessions by going around the group and hearing from each person what their biggest takeaway from the day was. In just a few words, what big or small thing has stuck with you?	
2. Let's begin with the first workshop session, in which the literature was presented, and then you engaged with the SEM. Did you experience any a-ha moments, where you felt you learnt something new or something clicked into place?	<ul style="list-style-type: none"> Did anything stand out in the feedback discussion? What insights did you take away from this session? Did the situation analysis help you think about something differently?
3. In the second session, you created an actor map. This was intended to help you identify actors in the space that you can collaborate or engage with. How worthwhile was this activity? Please elaborate.	<ul style="list-style-type: none"> Were the steps easy to follow? What insights did you take away from this activity?
4. After lunch, you discussed causal loop diagrams and CATWOW. How did you understand this?	<ul style="list-style-type: none">
5. Were there any ideas or concepts introduced in this session which you found confusing or unhelpful?	<ul style="list-style-type: none"> Did the language around and illustrations of the causal loop diagrams make sense? Did the soft systems methodology content (e.g., CATWOW) make sense?
6. In the final session of the workshop, you created the programme scope. What did you think of this session?	<ul style="list-style-type: none"> Was it useful? Did it feel like you were rehashing old ideas? Did the work from the prior sessions help to refresh these ideas?
7. How could today's process have been improved?	<ul style="list-style-type: none"> Clarity of concepts Time/pace Engagement with the group Flow Presentation format Resources/materials
8. Before we say goodbye, would anyone like to add something that hasn't come up yet?	

FG 2: Focusing Workshop- ToC

Set-up and Introduction

- Begin by thanking the team for participating in the workshop and for taking the time to provide feedback.
- Introduce the purpose of the focus group – to obtain feedback and suggestions for improving the workshop protocol. Stress that there are no right or wrong answers, and they need not hold back on constructive criticism.

Focus Group Questions

Questions	Probes
1. As we did last time, let's begin this session by hearing what everyone's biggest takeaway from today was.	
2. You began today with an introduction to theory of change, and then you went on to build your own theory of change. What did you think of the ToC-development process?	<ul style="list-style-type: none"> • Were the concepts clear/easy to understand? • How did you find the process of formulating outcomes and mechanisms? • What did you struggle with?
3. Did the development process shift your thinking in any way?	<ul style="list-style-type: none"> • Did anything new click into place? • Were there any a-ha moments? • Did it prompt you to frame something differently?
4. Suppose you had to take the ToC that you developed today and present it to a potential new funder or partner. Do you feel as though you would be able to explain it clearly and confidently?	<ul style="list-style-type: none"> • Would you be confident to explain it clearly? • Would an external stakeholder be able to understand your programme? • Would you want to change it, or simplify it first?
5. How did the work that was done in the Scoping / situation analysis workshop help with what you did today? Can you give any examples?	<ul style="list-style-type: none"> • How did the situation analysis help with developing the ToC? • What was particularly helpful?
6. How could today's process have been improved?	<ul style="list-style-type: none"> • Clarity of concepts • Time/pace • Engagement with the group • Flow • Presentation format • Resources/materials
7. Before we close the session, does anyone have any additional comments or reflections?	

FG 2: Focusing Workshop- ToA

Set-up and Introduction

- Begin by thanking the team for participating in the workshop and for taking the time to provide feedback.
- Introduce the purpose of the focus group – to obtain feedback and suggestions for improving the workshop protocol. Stress that there are no right or wrong answers, and they need not hold back on constructive criticism.

Focus Group Questions

Questions	Probes
1. As we did last time, let's begin this session by hearing what everyone's biggest takeaway from today was.	
2. To develop the ToA, you worked on various separate elements, including a participant definition, partner analysis, programme life space, and service utilisation diagram. What were your thoughts on these tools? Were any more useful or interesting than others?	<ul style="list-style-type: none"> • Were they clear/easy to understand? • What did you struggle with? • What was interesting? <p>Ask specifically about boundary choices, partner analysis, life spaces, service utilisation</p>
3. The ToA is meant to give you an in-depth and systemic view of your programme's implementation. To what extent do you think it achieves that purpose?	<ul style="list-style-type: none"> • How might you use the ToA for programme management? • How relevant is the ToA for your particular context?
4. How did the work that was done in the Scoping / situation analysis workshop help with what you did today?	<ul style="list-style-type: none"> • How did the situation analysis help with developing the ToC and ToA?
5. How has the timing between workshops been for you? Would you have liked them closer together or farther apart?	Ask about the time difference with the delay
6. How could today's process have been improved?	<p>Ask about doing the partner exercises etc themselves vs Hollie doing it for them on the flipchart/projector</p> <ul style="list-style-type: none"> • Clarity of concepts • Time/pace • Engagement with the group • Flow • Presentation format • Resources/materials
7. Before we close the session, does anyone have any additional comments or reflections?	

FG 3: Monitoring & Learning Workshop

Set-up and Introduction	
<ul style="list-style-type: none"> • Begin by thanking the team for participating in the workshop and for taking the time to provide feedback. • Introduce the purpose of the focus group – to obtain feedback and suggestions for improving the workshop protocol. Stress that there are no right or wrong answers, and they need not hold back on constructive criticism. 	
Focus Group Questions	
Questions	Probes
1. As usual, let's go around the room and mention one lesson or insight that stood out to you today.	
2. The first half of today's workshop revolved around developing a monitoring strategy to help you track your programme implementation and outcomes. How feasible do you think this strategy is to actually implement?	<ul style="list-style-type: none"> • Do you have a good grasp of the requirements? • Does your team have the requisite resources (including time and capacity)? • Do you feel adequately equipped to implement the strategy?
3. What kind of additional support, if any, do you think you will need to implement the monitoring strategy?	<ul style="list-style-type: none"> • Team capacity • Development of measurement tools • Fieldwork support • Training
4. In the second half of the workshop, you discussed organisational learning and put together a learning strategy. Some people might argue that learning takes place organically, and it cannot be planned for. What do you think?	<ul style="list-style-type: none"> • Does the strategy that you developed have relevance and utility for your work? • Do you think the strategy will be helpful? • Do you think the mechanisms you put in place today will support your learning as a team?
5. How do you envision the learning components that we discussed today fitting in, or not fitting in, with your established routines and schedules (e.g., learning events, data analysis and reporting)?	<ul style="list-style-type: none"> • What changes will you be making based on today's activities? • Is there room in your busy schedules and your minds for the kinds of activities discussed today? • Does this complement your work, or add additional burden?
6. How could today's process have been improved?	<ul style="list-style-type: none"> • Clarity of concepts • Time/pace • Engagement with the group • Flow • Presentation format • Resources/materials
7. Before we close the session, does anyone have any additional comments or reflections?	

FG 4: Evaluation Workshop

Set-up and Introduction	
<ul style="list-style-type: none"> Begin by thanking the team for participating in the final workshop and for taking the time to provide feedback in all of the focus groups sessions. Introduce the purpose of the focus group – to obtain feedback and suggestions for improving the workshop protocol. Stress that there are no right or wrong answers, and they need not hold back on constructive criticism. 	
Focus Group Questions	
Questions	Probes
1. For the last time, let's go around the room and mention one lesson or insight that stood out to you today.	
2. The first half of today's workshop involved designing a future evaluation. How did you find this session? Was it easy to follow or challenging - why?	<ul style="list-style-type: none"> Was it easy to follow and understand? Were any ideas or concepts challenging? Was anything irrelevant?
3. Following design, you moved on to making key boundary decisions for an evaluation. What are your thoughts on this session?	<ul style="list-style-type: none"> Was it useful? Was it relevant? Was it interesting? Was it challenging? Was it boring?
4. Many of you may not have any prior experience or knowledge of evaluation. Based on the work done today, how prepared do you feel for evaluation?	<ul style="list-style-type: none"> How did the workshop affect assumptions that you may have had on evaluation? Did it help you understand more about evaluation?
5. What did you think of the evaluation game that you played?	<ul style="list-style-type: none"> Was it interesting or a waste of time? What did you learn from it?
6. The workshop closed today with a team debrief on evaluation readiness. What insights from this session stood out to you?	<ul style="list-style-type: none"> What concerns were raised? What support mechanisms do you think are needed?
7. How could today's process have been improved?	<ul style="list-style-type: none"> Clarity of concepts Time/pace Engagement with the group Flow Presentation format Resources/materials
8. Today was your final workshop and this is our last focus group. Can we take a moment each to provide any final reflections on this process. Let's start by going around the room and each person naming one thing that they learnt about that they are taking away from the process? This can be an idea or an activity that you did. Any final reflections or comments on the process? How do you envision taking this forward?	<ul style="list-style-type: none"> Examples of concepts or activities: BATWOE; programme theory; life spaces; service utilisation diagrams; 3-R model, learning events, critical systems heuristics; context What is your biggest takeaway from all of the workshops?
9. Before we close the session, does anyone have any additional comments or reflections?	

Appendix I: Observation Schedules

Organisation: _____ Time: _____
 Date: _____ Meeting details: _____
 Place: _____ Staff in attendance: _____

	Observations	Reflections
Physical Setting <i>Please describe the layout of the space, how participants are seated, and what equipment and resources are displayed and used.</i>		
Session 1		
Workshop Activities & Time <i>List the order of events, and the time taken to complete each one.</i>		
Participant Engagement <i>Probes: Interest, attention, questions, confusion, understanding, conversation.</i>		
Facilitator <i>Probes: Pace, explanation, engagement, respect, inclusivity.</i>		
Session 2		
Workshop Activities & Time <i>List the order of events, and the time taken to complete each one.</i>		
Participant Engagement <i>Probes: Interest, attention, questions, confusion, understanding, conversation.</i>		
Facilitator <i>Probes: Pace, explanation, engagement, respect, inclusivity.</i>		

Session 3

Workshop Activities & Time <i>List the order of events, and the time taken to complete each one.</i>		
Participant Engagement <i>Probes: Interest, attention, questions, confusion, understanding, conversation.</i>		
Facilitator <i>Probes: Pace, explanation, engagement, respect, inclusivity.</i>		
OTHER NOTES		

Appendix J: Final Debriefing Interview Schedule- Heather

1. As you know, the toolkit is meant to instill an understanding and appreciation for systems thinking among evaluators. Would you say this was achieved to any extent with you? Do you feel like you understand systems thinking?
2. What were your major learnings from this process?
3. Is there anything you would like to take forward into your own work?
Where did you struggle the most?
4. Walk me through your process – how did you engage with the chapters and then weave them into your workshops?
5. Did you enjoy reading the chapters' content?
6. I'm going to name each chapter – what are like one or two key words, ideas, or concepts that appear in your mind when I name them?
7. What did you think of the worksheets? Were they a helpful addition or more of a burden?
8. Let's talk about the actual workshops now. How did they differ to what you normally do?
9. Again, if we mention each one, what stands out in your memory per workshop?
10. As a random evaluator who was not approached for this pilot, how interested would you be in a book or product or whatever that takes you through this process?
11. What are your recommendations for taking this process forward?

Appendix K: Follow-up Interview Schedule – NPO Directors

Set-up and Introduction	
<ul style="list-style-type: none"> • Begin by thanking the interviewee for taking the time to provide feedback in this session. • Introduce the purpose of the interview – to understand the extent to which the framework has been put to use in the organisation, and whether it has helped or hindered the team’s work. Stress that there are no right or wrong answers, and they need not hold back on constructive criticism. 	
Interview Questions	
Questions	Probes
Can you please describe the immediate aftermath of the workshops? Were there any further discussions among the team, or particular actions that were taken?	<p>Were additional meetings held to discuss strategy and planning?</p> <p>Were teams assigned new tasks based on the workshop outputs?</p> <p>Were new activities planned?</p>
Let’s talk specifically about the work that we did in the workshops, and what your intentions around these outputs are now. Have you done anything with the programme scope in the last month? Or do you intend to?	<p>E.g., communicated it via your website, social media, annual reports, used it in proposals.</p> <p>Have you refined it further?</p>
We used the situation analysis findings in the course of the workshops. Has it served any additional uses since the workshop?	E.g., used for proposals, communicated with others
What impact did the programme theory have on programme planning and strategy? Do you think differently about anything now?	<p>E.g., targeting, implementation planning, programme content, scheduling, staffing</p> <p>Has the programme theory changed the way you think about your programme content, implementation, context or strategy?</p> <p>Are you planning to make any adjustments to your strategy based on the programme theory?</p>
How do you intend to use the programme theory going forward?	E.g., programme planning, evaluation planning, monitoring, communications
After we developed the monitoring plan, has any further work taken place to realise the plan?	E.g., staff allocation, scheduling, instruments, data collection, tools, etc.
What, if any, additional support do you feel is needed to put the monitoring plan into action?	E.g., finances, capacity, advice, fieldworkers
What steps have been taken in line with the learning strategy? Have any plans been put into place?	<p>Have you planned any learning events?</p> <p>What has the learning champion done since the workshop?</p>
Are there any general concerns about any aspect of the MEL framework going forward?	E.g., alignment with reporting requirements, workload, finances, capacity, time
One of the major goals of the SAMEL toolkit is to help develop and encourage a learning culture in organisations. To what extent do you think this has happened at your organisation?	
As you know, the toolkit draws on systems thinking approaches and tools. How did the workshops help you learn about the system in which you work?	
Do you have anything more to add?	

Appendix L: Final Codes

Theme 1: Strategic clarity: Harnessing systems thinking for programme theory				
Codes	Files	References	Code Description	Illustrative Quotes
Identifying opportunities	4	6	Clarity in the identification of potential opportunities that are available to the organisation, or will become available, such as funding, partnerships or future programmes.	<p>“And if you look at that value chain of all the things that we were talking about, it makes you think of the next thing, what is the next thing? How do we remain one step ahead of everyone, because EYEN is one of the trendsetters in the sector. So you know, that process for me, it just made me think, what is the next thing that we need to focus on here now, but how do you push forward?” [fg]</p> <p>“I think that gives us this other conversation. It gives us good business ideas and other models for things that we can work on developing, like Unathi was saying- the funding, additional funding, whatsitcalled? Subsidies. And there's those opportunities. Ya so it was good to see like the other opportunities that kind of, maybe, at the end of the rainbow.” [fg]</p>
Impact focus	7	12	Clarity in the social change outcomes and impact that the organisation wants to achieve among their communities. This includes drawing conceptual boundaries around the organisation’s envisioned contribution to change (impact and outcomes) and the discussions that define those boundaries.	<p>“I think the most useful part for me was just the one simple question- is 'where do you want to get to?' And so we were able to cut through everything. And say that, like 'do 100 things, does that take us there?' “. [fg]</p> <p>“I like Jessica's way of facilitation, because at most when we write proposals and we write things you sort of copy a document elsewhere. Now it came [unclear] start thinking wena, now what we are doing? [laughs]...Because now think about it- like now- what do you intend to achieve? What do you intend to change? You see, those kind of things. But also, what would you see as effects of the change that you've started here and to the future.” [fg]</p>
Implementation barriers	3	4	Clarity around the barriers to effective programme implementation/delivery that are caused by the organisation itself.	<p>“The barriers, because it gives you really key concrete places to intervene and fix what you're doing in a very straightforward way that I think a lot of people might not even consider, because it always feels really overwhelming. When you look at it, like, 'oh, it's communication...we can fix that'.” [fg]</p>
Intervention bounding	15	35	A clarified process of deciding what boundaries to draw, and how, around the programme’s interventions to meet community needs in an effective way.	<p>“So it also helps with that clarification, because sometimes you see all these things, and you have question marks, and it stresses you out because like we should be doing that, we should be doing this. But this just says we don't need to. It helps to reinforce and it actually helps with focus. Because there's a million things we can do. But we don't need to do all of them.” [fg]</p> <p>“Centre support & registration support turn out to be different programmes and there is a collective meltdown. Zara walks out, comes back smelling like smoke. Kate exclaims: ‘The next training we need is media training because we cannot face the world like this. We are a mess!’ Lots of shouting and laughing. Heather reigns things in by saying ‘this needs to be a conversation that needs to be had internally to determine the boundaries of this programme’.” [obs]</p>
Planning	4	7	Clarity around future plans and preparedness, in terms of programme delivery and evaluations.	<p>“So I'm just saying that now I'm putting this into action, visualizing it out, it will come up. And of course, this will produce a first class business plan that needs to be costed and all sorts. So I'm seeing this as a valuable experience.” [fg]</p>

				<p>“And I think it's something that traditionally is easily overseen. I mean, you quickly write up the terms of reference, and poof, off it goes. But to have broken it down in this way is kind of like, detailed, and it does exactly that. It gives it a very, very good structure, in a sense, so you've already pre planned what happens in the event of, you know.” [fg]</p>
Relationships with stakeholders	5	14	Clarity around the relationships the organisation has or wants to have with its external stakeholders and programme partners.	<p>“And also our role in that relationship...To become a lot more clear. 'This is government's role. We don't want to take that on. How do we work with government, and position ourselves in that context?’” [fg]</p> <p>“We kind of refined the way we think about what we are capable of doing as the organisation, and what we can commit to, as an organization for delivery. And then also identifying who those partners are that should be working hand in hand with us in order to achieve the greater national or country goals, you know.” [org debrief interview]</p>
Shared programmatic understanding	13	26	A collective sense of ‘being on the same page’ among the NPO team about the programme under focus, in terms of understanding what the programme does and what it intends to achieve.	<p>“And the process with you allowed me to understand the nitty gritty of the implementation of the program much better. So now, when Pumeza speaks about stuff, I understand, and I can connect with the things much clearer and better, and give more strategic support and advice now, because I have a better understanding of it now” [org debrief interview]</p> <p>“I think it really brought everybody onto the same page. It allowed us to all have a similar understanding of what the program is meant to do, and how we are envisioning it to reach the outcomes we want it to reach.” [org debrief interview]</p>
Strategic clarity	1	2	Clarity around the broader organisational strategy.	<p>“For me the theory of change is going to go to link to strategy. And I mean, we've been here [unclear]...reviews. And strategy is extremely complex, and particularly this aspect of strategy. And I just loved how, in a very short space of time, we were able to crystallize and synthesize the complexity.” [fg]</p>
Programme design	12	32	The strategic/conceptual design of the programme and its interventions by the management team.	<p>“So the program itself is there's different component parts to the program, and it's not been first merged into one holistic design. And so we're also still busy with the program design. But the importance of the framework though- it gives us the framework in terms of which to create the design, as per the outcomes and theory of change.” [fg]</p> <p>“That, I think, is critical for us to relook at program design. Because now if you can see where those drop off points are, you know where your system strengthening efforts needs to be targeted at.” [fg]</p> <p>“Gail comments that evaluation questions should be asked during programme design – upfront; ‘It’s almost like we don’t know what the outcomes are going to be – they’re a surprise! It shouldn’t be like that!’ “. [obs]</p>

Theme 2: Contextual disconnect: Bridging theory and the complex realities of non-profit environments

Codes	Files	References	Code Description	Illustrative Quotes
Accountability to community	5	8	A feeling of duty and accountability experienced by NPOs toward their community, particularly in terms of providing a contextually appropriate solution to a problem identified in collaboration with the community.	"But also, I think, for me, it is who was involved in the situational analysis, because now we need to have an information session for the whole- for the principals. So now, when we have that information session, we'll share a plan, informed by their input during the situational analysis. To say that- this is our response then, to your situation, and what Eduspark commits to implement." [fg]
Donor influences	12	29	The power and influence that the NPOs' donors have over the organisation's MEL design and implementation, such as the reporting requirements from donors that dictate the data that must be collected or the evaluation questions that must be asked.	"And a lot of NGOs get dented because of that donor pressure comes in. And I think we've probably suffered from a very, you know, in a hard way, even now with M&E sections being fragmented, because that's what the donor wanted." [fg] "We understand the nature of the reports and having to deal with thirty different funders on a weekly basis. That's the info they want to see. And unfortunately, that's what we need to provide so and so this this will help scope it better." [fg]
Influence of ECD policy	9	10	The influence that the requirements of the Western Cape ECD norms and standards have on the organisations' programme design and implementation plans (and thus MEL design).	"Quite a bit of debate around whether centre registration should be a defining criteria for quality and other evaluation criteria." [obs] "Mandla raises the issue of health and nutrition, Heather asks if it's in the interventions and indicates that the ToC is what EYEN is committing to. Mandla explains that this is part of compliance." [obs]
Programme design contentions post Focusing	5	9	Ongoing debates and discussions around programme design and theory elements that took place in the M&L and Evaluation workshops.	"Because internally, even up until like this last question... people still have very different ideas of what EYEN's strategic goals and visions are, and how it's tied to the program, and how we execute those." [fg] "Centre support & registration support turn out to be different programmes and there is a collective meltdown." [obs]
Programme design work post pilot	2	5	Changes to the programme design mentioned to take place after the SAMEL pilot had concluded.	"I mean, we want what we're doing across all the elements of our organization to align with what the organisational strategy is moving forward- not only for this year, but for the next five years, for example. So we want to streamline it so that it can cohesively work together with each other. Not necessarily changing or adapting things to a degree where it's completely different to what it was before, but trying to find the synergies between the two." [follow-up interview]
The overwhelm of complexity	7	9	The feeling of overwhelm among the participants and evaluators when thinking about, discussing, or mapping the complex social and environmental context in which the organisations operate.	"And so it's just rather thinking about it like that, because it's daunting. It's extremely daunting, it's extremely overwhelming. It almost feels like , you know, 'I don't think it's gonna work, let's just throw in the towel, let's just run away'. Because that's quite often how you get to feel- it's too much. But I think that, you know, you maybe just also sometimes need to mind your business, and stay in your lane, and influence what you can." [fg] "Heather is very overwhelmed by the messiness of the interventions as she tries to neaten it up during lunch. She asks: 'Is this what ECD is like?'" [obs]

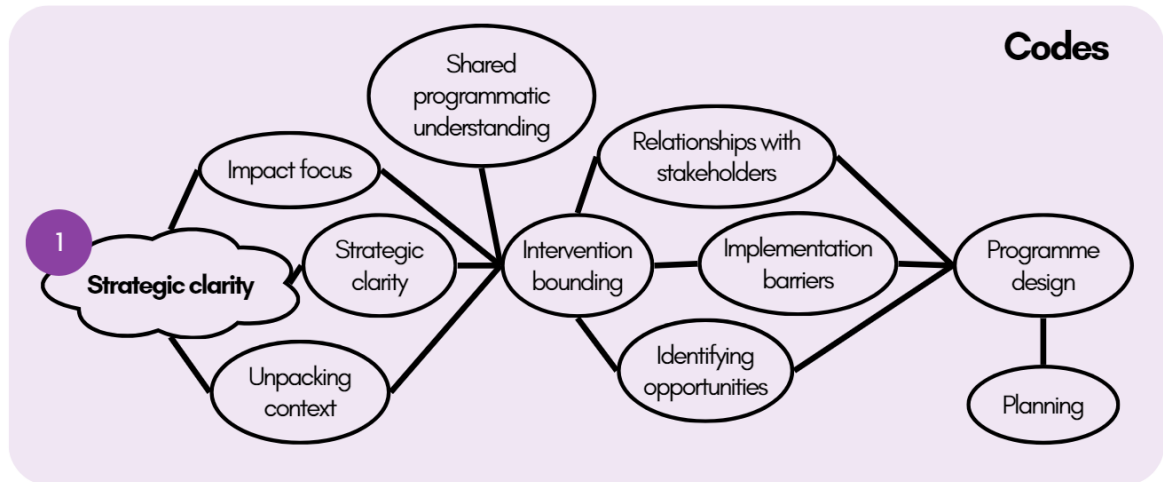
What's ideal vs realistic	5	6	A strategic conflict applicable to drawing boundaries around MEL and programme elements, such as targets or measurement areas. Seen in participant debates around what is ideal versus what is realistic.	<p>“A: Or, are we not dreaming too much? B: She's always worried about the size and the scope of it. And the reach.” [fg]</p> <p>“We're looking at this thing from a holistic kind of implementation perspective. So we can't narrow, we need to look at the full spectrum, if you want to touch and do this holistic approach, which I think started out the programme. So we need to map out the <i>whole</i> thing and so on, and see where we are at. And then obviously, this is going to lead us to what is manageable and what is not..But we need to do it in that full scope. So that we see because this is our vision. We can't have that vision and then say 'but we can't do this no no’” [fg]</p> <p>“Mandla says with a cynical smile : ‘ideal situations make life so much easier’” [obs]</p>
---------------------------	---	---	---	---

Theme 3: Organisational disconnect: Bridging MEL development with organisational readiness for MEL				
Codes	Files	References	Code Description	Illustrative Quotes
Confusion in vision & mission		8	Poor strategic clarity and/or organisational alignment on the organisation's vision and mission.	<p>"Because internally, even up until like this last question... people still have very different ideas of what EYEN's strategic goals and visions are, and how it's tied to the programme, and how we execute those. So for me, I would love to have a conversation afterwards" [fg]</p> <p>"We're sitting with information, we've got it written out. Bits there, bits there, some in more detail than others, and a lot in our heads. And we haven't gone through a structured process of unpacking all of that and processing all of that information- ideas, thoughts, vision.." [fg]</p>
<i>MEL Constraints</i>				
Organisational buy-in		4	The shared and agreed-upon collective responsibility for MEL among all organisation members, including an understanding of the MEL framework.	"So I think the concern I have is just the understanding and role of M&E across the organization. And that, you know, we need to find a way where people realize that everybody's responsible for this. And everybody's got a role to play in delivering and using this framework" [fg]
Organisational capacity for MEL		12	The personnel capacity, and lack thereof, that is available to conduct MEL-related tasks and activities.	<p>"I'm worried about capacity and, you know, people's roles and assigning roles and expecting people to do that in the current constraints of us not having this and a that and a that post, and so on." [fg]</p> <p>"But if we continue the way we are on a paper based method, there's no way in hell we have enough capacity for that." [fg]</p>
Quality and training		4	The quality of programme data that is collected in the field, and the associated training needs required to increase quality.	"Looking at the collection tools. So we've identified certain collection tools, but if you now look at each and every one of those tools... how do we... I think what scares me now is looking at integrating those tools into something that's simple enough that you can actually tick and get what you need, with a signature, without going with a wad of paper." [fg]
Tools		6	Data collection instruments used to collect MEL data in the field.	<p>"I was even thinking, deciding on the tools to be used for evaluation, it's not gonna be easy. Because there are areas where we can't even take our devices to. We can't. We need to think of hard copies and stuff, because you can't take a laptop, you can't take your phone, you can't take anything valuable." [fg]</p> <p>"Looking at the collection tools. So we've identified certain collection tools, but if you now look at each and every one of those tools... how do we... I think what scares me now is looking at integrating those tools into something that's simple enough that you can actually tick and get what you need, with a signature, without going with a wad of paper." [fg]</p>
Lack of learning culture		14	The total lack or poor learning culture in the organisation, evidenced by not institutionalising learning resources and policies, low levels of behavioural norms around learning, and a lack of personnel capacity to hold and implement a learning system.	"I think that the concern for me, is often where M&E gets held. And maybe I shouldn't even call it M&E, we should call it the continuous, you know, learning, because sometimes I just leave it in the hands of one or two people. So it's either the M & E person, and they expected to drive it and it's seen as their role. And so when things get crazy or you're caught up and the role doesn't get fulfilled then that's the department we blame, because it's not happening. So I think the concern I have is just the understanding and role of M&E across the organization."

Programmatic siloes		8	An organisational approach to the management of separate programmes in which staff, work and strategy are siloed per programme.	<p>“And really the thinking is that we've got five different programs in this organization. And as an organization, it's hard. Because it's pulling your resources and competing with different things. But we need to almost look at it as if we were five different organizations and actually ringfence, and support, and actually have those teams. Because me as a CEO, I can't run five organizations, but we can have five CEOs run five organisations.” [fg]</p> <p>“And I think also it comes from the point where, I know [external partner] is very structured. And they have their elements already built in, and there's elements were the data collection and the submissions that you need to do at a certain time, obviously, is built into the programme structure. So I think that's why there's also not been much crosstalk between my department and your department purely because there's a structure in place from an external point of view, right?” [fg]</p>
Team investment & commitment		7	The amount of time, energy and capacity that the organisations invested into the pilot process/workshops.	<p>“You know, particularly now, as it's crunch time for us in terms of where we are currently. Do we have the time to invest in this? Can we afford time out of the office, a whole day in this case? Will it be of any benefit to us? And I can I can say definitely, for me personally, as a CEO of organization, there's a lot of things that- information I took- or things I became aware of that... I made side notes myself to follow up on and to pursue” [fg]</p> <p>“I feel slightly anxious about the idea of repeating this again for all the other programs. [laughter].” [fg]</p>
The luxury of strategising vs operations	0	14	The need of the organisations to prioritise time toward programme operations, as opposed to coming together in teams to work on high-level strategy.	<p>“But unfortunately, you've got operational responsibilities, which just trump all of this luxury of having to put together these processes and system and time. So we have to bring capacity in to take over the operational support, so you can start putting those systems in place, but it doesn't happen overnight.” [fg]</p> <p>“But I do think that management needs to spend more time together. I felt like this was the one of the few occasions we get everyone in the room where we are be able to, like, have a little bit of fun, get to see a little bit of everyone's personalities. And we don't often have the opportunity to do that.” [fg]</p>
Team discussion & debate	6	38	Instances and opportunities in the workshops for collective, constructive discussion and meaningful debate among the participants about the programme design and the design of the MEL framework.	<p>“We move on to curriculum support and this takes forever because there is still so much disagreement or non-consensus around what the programme is and will do. We debate for ages around ‘clusters’ – Timothy and Zola have totally different definitions of what this is and how it fits into the programme offering.” [diary]</p> <p>“It's just for me, once again, I mean, the discussions here just stimulate thinking in my head around the bigger picture, the organizational stuff, and so on. So that, to me is quite an important concern. I get somehow frustrated by the two of them becoming so emotional on the heated stuff [laughter].” [fg]</p>

Appendix M: Theme Descriptions

Theme 1: Strategic clarity: Harnessing systems thinking for programme theory



Theme Description:

This theme encapsulates the strategic advantages that resulted from the SAMEL process for the NPO teams. It captures the clarity that emanated from the Scoping and Focusing workshops in terms of the programme theory. The insights gained from theory development feed into more effective programme planning.

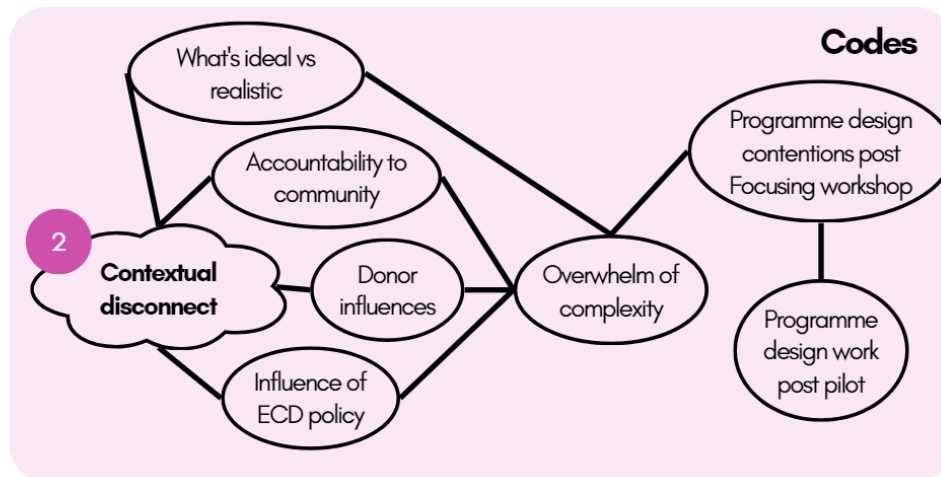
Key Findings:

Scoping, ToC and ToA workshops were incredibly valuable to teams in terms of strategic programme design and planning.

The workshops helped teams gain a lot of clarity around programme goals and intervention design.

Systems thinking concepts and tools seemed to enhance the critical thinking process.

Theme 2: Contextual disconnect: Bridging theory and the complex realities of non-profit environments



Theme Description:

This theme encapsulates the ongoing tension between an academic process and its practical application in resource-limited settings. Non-profit work involves providing much-needed services with limited resources within complex social contexts, characterised by interactions between environmental constraints, policy pressure, and donor influence. This can lead to a feeling of overwhelm among NPO teams. This complexity can affect the development of programme theory, as team debates and discussions around programme design can be lengthy and require time to resolve before the MEL framework can be developed effectively.

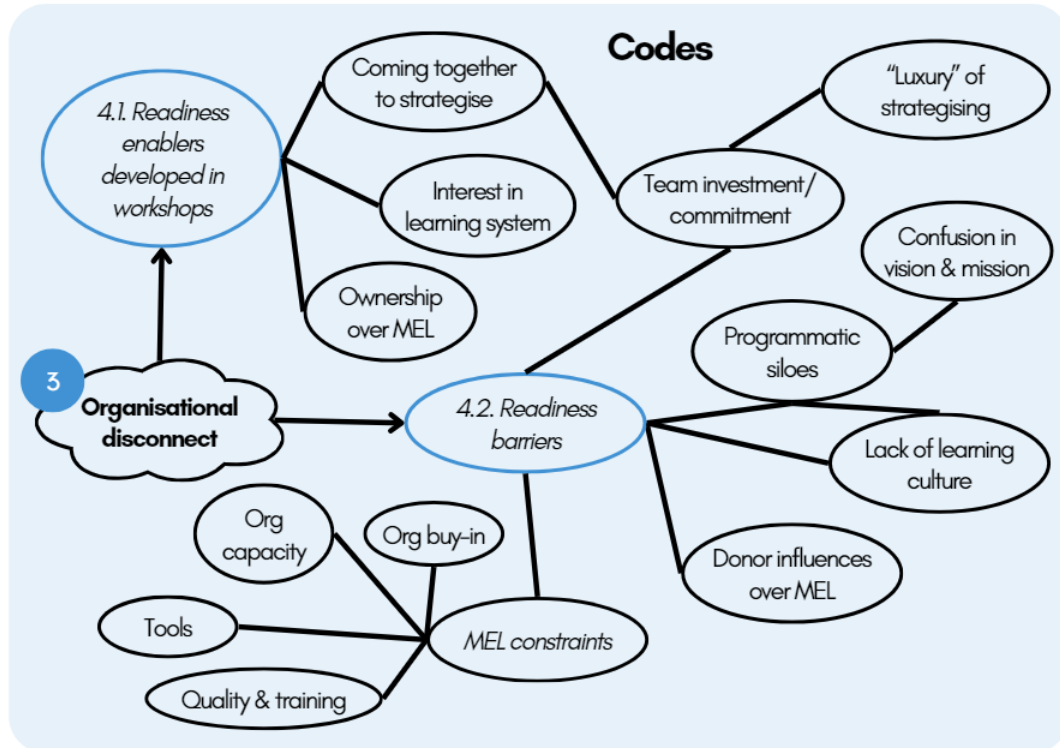
Key Findings:

Programme theory development was complicated by NPO sources of complexity: policy, donors, resource constraints, community need.

Discussion/debate around programme design extended beyond the Focusing workshops, and even after the pilot.

This limited the ability to properly engage with MEL design.

Theme 3: Organisational disconnect: Bridging MEL development with organisational readiness for MEL



Theme Description:

This theme highlights the organisational features that emerge out of complex sectoral contexts and may constrain MEL processes, specifically illustrating how the resources, capacity, structure and culture of the organisation play a major role in determining MEL readiness, and thus, the SAMEL workshop outputs and outcomes.

Key Findings:

Due to the complex contexts that they work in, the teams were not MEL-ready and lacked a learning culture as operations are prioritised above strategic reflection, and MEL is seen primarily as an accountability tool.

The workshops successfully developed some MEL readiness enablers (bringing the team together, developing a sense of ownership, developing interest in learning) but more work was needed to overcome barriers to MEL and learning.

Appendix N: Ethical Approval Letter



Faculty of Commerce

Private Bag X3, Rondebosch, 7701
2.26 Leslie Commerce Building, Upper Campus
Tel: +27 (0) 21 650 4375/ 5748 Fax: +27 (0) 21 650 4389
E-mail: jacques.rousseau@uct.ac.za
Internet: www.uct.ac.za



@Commerce UCT



UCT Commerce Faculty Office

26 08 2021

Jessica Horler
School of Management Studies
University of Cape Town
REF: REC 2021/08/023

**A systems-based approach to strategic planning, monitoring and evaluation:
Designing and testing the SAM&E Protocol**

We are pleased to inform you that your ethics application has been approved. Unless otherwise specified this ethical clearance is valid until 31-Dec-2022 .

Your clearance may be renewed upon application.

Please be aware that you need to notify the Ethics Committee immediately should any aspect of your study regarding the engagement with participants as approved in this application, change. This may include aspects such as changes to the research design, questionnaires, or choice of participants.

The ongoing ethical conduct throughout the duration of the study remains the responsibility of the principal investigator.

We wish you well for your research.

2021.08.26
12:20:05 +02'00'

Jacques Rousseau
Commerce Research Ethics Chair
University of Cape Town
Commerce Faculty Office
Room 2.26 | Leslie Commerce Building

Office Telephone: +27 (0)21 650 2695 / 4375
Office Fax: +27 (0)21 650 4369
E-mail: jacques.rousseau@uct.ac.za
Website: <http://www.commerce.uct.ac.za/com/Ethics-in-Research>

"Our Mission is to be an outstanding teaching and research university, educating for life and addressing the challenges facing our society."

Approval of Extension



Faculty of Commerce

Private Bag X3, Rondebosch, 7701
2.28 Leslie Commerce Building, Upper Campus
Tel: +27 (0) 21 650 4375/ 5748 Fax: +27 (0) 21 650 4369
E-mail: jacques.rousseau@uct.ac.za
Internet: www.uct.ac.za



@Commerce UCT



UCT Commerce Faculty Office

11 10 2022

Jessica Horler
School of Management Studies
University of Cape Town
REF: REC 2021/08/023

**A systems-based approach to strategic planning, monitoring and evaluation:
Designing and testing the SAM&E Protocol**

We are pleased to inform you that your ethics approval has been extended. Unless otherwise specified this ethical clearance is valid until 31-Dec-2023 .

Your clearance may be renewed upon application.

Please be aware that you need to notify the Ethics Committee immediately should any aspect of your study regarding the engagement with participants as approved in this application, change. This may include aspects such as changes to the research design, questionnaires, or choice of participants.

The ongoing ethical conduct throughout the duration of the study remains the responsibility of the principal investigator.

We wish you well for your research.

A handwritten signature in black ink, appearing to read 'JRousseau'.

2022.10.11
10:51:17 +02'00'

Jacques Rousseau
Commerce Research Ethics Chair
University of Cape Town
Commerce Faculty Office
Room 2.26 | Leslie Commerce Building

Office Telephone: +27 (0)21 650 2695 / 4375
Office Fax: +27 (0)21 650 4369
E-mail: jacques.rousseau@uct.ac.za
Website: <http://www.commerce.uct.ac.za/com/Ethics-in-Research>

"Our Mission is to be an outstanding teaching and research university, educating for life and addressing the challenges facing our society."

Appendix O: Data Management Plan

A systems-based approach to strategic planning, monitoring and evaluation: Designing and testing the SAM&E Protocol. - Student Outline DMP

1. General guidelines

PURPOSE OF THIS TEMPLATE - The purpose of the Outline DMP is to indicate your initial plans for how your data will be collected, shared and stored, and to give you a chance to think about these data-focused aspects of the research process. As you begin doing your research, your data process may change, and it is perfectly acceptable to change your data management plan to accommodate the changes in your research process. Indicate below that you understand the purpose of completing this Outline DMP template.

- I understand the Outline DMP template is a projection of my anticipated data management planning requirements and should be updated as my project develops.

2. Authors and supervisors

PROJECT NAME - Replicate the title of your project, dissertation or thesis exactly as it appears in your proposal document.

A systems-based approach to strategic planning, monitoring and evaluation: Designing and testing the SAM&E Protocol

PERSONAL DETAILS - Indicate the name(s) and student number(s) of the student(s) who will be involved in this project, dissertation or thesis.

Jessica Horler
HRLJES001

SUPERVISOR(S) DETAILS - Indicate who will supervise this project, dissertation or thesis. If you do not yet have a supervisor, leave this section blank.

Prof Sarah Chapman
Dr Carren Duffy

3. Data Collection/Generation

COLLECTION OF ORIGINAL DATA - Indicate whether or not you intend to gather/produce original data for your study, and provide a brief description of the kind of data you think you will collect. If you are unsure at this time, indicate what you think you are most likely to collect. If you are not intending to gather or collect your own data, declare that here.

- I intend to collect original data (described below).

I intend to collect both quantitative and qualitative data. The qualitative data will take the form of three focus group interviews (final sample is yet to be determined), and three one-on-one interviews. The data will be transcribed in MS Word, and transferred to N.Vivo for analysis. I anticipate the dataset will be between 100MB and 300MB.

The quantitative data will take the form of two Google Form surveys. One is an eligibility form that will be distributed across my networks widely for the purpose of case recruitment. Responses will be transferred to MS Excel. I anticipate this dataset will be less than 100MB. The second survey will be distributed to the personnel of three organisations (final sample is yet to be determined). This data will also be transferred to MS Excel. I anticipate this dataset will be less than 100MB.

USE OF EXISTING DATA - Indicate if you intend to re-use existing data, either from online searches or from datasets provided by your supervisor, lab, or funder. If you are not intending to re-use existing data, declare that here.

- * I intend to reuse existing data in my study (described below).

I will request and review programme documents from the NGOs that make up my case studies. This includes whatever existing records they are willing to share with me, such as annual reports, evaluation reports, programme histories, etc.

DATA SHARING - Indicate whether or not you are intending to publish your research data. If you are, indicate where you are intending to publish your data and under what licensing conditions, such as Creative Commons. If you are not intending to publish your data, provide reasons and reference the appropriate ethical considerations, commercial applications/patenting ambition, or data re-use agreements that prevent you from publishing your data.

- * I do not intend to share my data because of confidentiality issues.

I do not intend to share my data as it is qualitative in nature and specific to local NGOs. It may be difficult to disguise these NGOs in the datasets, which would be in breach of my ethical responsibility to respect their privacy and anonymity.

4. Data Storage

ANTICIPATED DATASET SIZE - Indicate the estimated size of your completed dataset, and indicate whether or not you will need to access additional data storage facilities. If such storage is not provided by your unit or department, you may need to factor in the cost of purchasing additional storage space.

- * 20GB or less

DATA BACKUPS - Indicate how you plan to ensure your data is secure and retrievable in case of errors or hardware failure. Describe what procedures you will put in place to back-up copies of your data and where they will be stored.

- * I intend to backup my data using a service provided by UCT (UCT GoogleDrive, UCT OneDrive, Netstorage, ZivaHub etc.).

5. Data Centre(s)/Repositories

DATA CENTRES/REPOSITORIES - Once your project, dissertation or thesis is complete, it is advisable to curate and archive your completed dataset with an established data centre or repository. Note that you should archive your data even if you are not intending to publish it. Check with your supervisor or funder if you are required to deposit your data in a specific repository, or declare that you will deposit the data in ZivaHub (see the Guidance section).

- * At the end of my study, I will deposit my data on ZivaHub.

METADATA - Metadata is descriptive information that others will need to make sense of your dataset. Metadata includes things like study descriptions or abstracts, study instruments (sample collection schedules, codebooks for variables, survey instruments, etc.), subject codes, and keywords. Indicate what metadata will accompany your curated dataset.

The dataset will be accompanied by a short description from my dissertation abstract and relevant information from the methods section. It will also contain relevant codebooks.

6. Budget

BUDGET - Indicate any costs specifically relating to the management and curation of your data, such as purchasing additional storage space, digitisation of physical media, data storage or curation charges, and data audits. Most student research will be able to make use of free options provided by UCT and will not have to budget for data costs.

- * I do not anticipate any data costs as my data is less than 10GB, and I will be using a storage system provided by UCT (UCT GoogleDrive, UCT OneDrive, Netstorage, ZivaHub, etc.) to curate my data.

Appendix P: Workshop Participant Consent Forms

NPO Programme Team



Consent Form: Programme Team

Study: A systems-based approach to strategic planning, monitoring and evaluation in ECD: Designing and testing the SAMEL Toolkit
Researcher: Jessica Horler



You have been invited to participate in a doctoral research project conducted by Jessica Horler from the University of Cape Town. This project will design and test a systems-based toolkit for the development of a monitoring, evaluation and learning (MEL) framework for use by nonprofit organisations in the Early Childhood Development sector.

Your organisation has been selected to take part in the pilot testing of this toolkit. The objective of this process is to run through the toolkit via five workshops. During these workshops, you and your team will collaborate with the researcher to develop a MEL framework for your programme.

Your participation in this pilot involves participating in all of the workshops. You will also be asked provide feedback regarding your experience in a focus groups after each of the workshops (with the rest of your team).

Please take note of the following:

- Your participation in this pilot is voluntary.
- You have the right to withdraw from the pilot at any time.
- The workshops and focus groups will be audio-recorded and transcribed.
- You, and your organisation, will **not** be identified by name in the dissertation report, nor any associated reports.
- Information that you provide as part of this research will be subject to standard data use policies which protect anonymity of individuals and institutions.

Please indicate your consent for this research by responding to the conditions below:

- | | |
|--|---|
| <input type="checkbox"/> I understand the purpose of the research | <input type="checkbox"/> I understand that my feedback will be recorded so that the researcher may accurately portray my responses |
| <input type="checkbox"/> I understand that I can withdraw from this research at any time without giving any reason | <input type="checkbox"/> I understand that my identity (and that of my organisation) will remain confidential and, where appropriate, a pseudonym will be used in referring to information that I provide |
| <input type="checkbox"/> I understand that the information that I provide may be included in an institutional report | <input type="checkbox"/> I consent to take part in this research. |

Name
First Name Middle Name Last Name

Signature Date
MM DD YY

Programme Stakeholders

 **Consent Form: Programme Stakeholders**
Study: A systems-based approach to strategic planning, monitoring and evaluation in ECD: Designing and testing the SAMEL Toolkit
Researcher: Jessica Horler



You have been invited to participate in a doctoral research project conducted by Jessica Horler from the University of Cape Town. This project will design and test a systems-based toolkit for the development of a monitoring, evaluation and learning (MEL) framework for use by nonprofit organisations in the Early Childhood Development sector.

The organisation that you work (EduSpark) with has been selected to take part in the pilot testing of this toolkit. The objective of this process is to run through the toolkit via five workshops. During these workshops, the team will collaborate with the researcher to develop a MEL framework for your programme.

Your participation in this pilot involves participating in one or some of the workshops. You may be asked to provide feedback regarding your experience in a focus group after each of the workshops.

Please take note of the following:

- Your participation in this pilot is voluntary.
- You have the right to withdraw from the pilot at any time.
- The workshops and focus groups will be audio-recorded and transcribed.
- You, and your organisation, will **not** be identified by name in the dissertation report, nor any associated reports.
- Information that you provide as part of this research will be subject to standard data use policies which protect anonymity of individuals and institutions.

Please indicate your consent for this research by responding to the conditions below:

- | | |
|--|---|
| <input type="checkbox"/> I understand the purpose of the research | <input type="checkbox"/> I understand that my feedback will be recorded so that the researcher may accurately portray my responses |
| <input type="checkbox"/> I understand that I can withdraw from this research at any time without giving any reason | <input type="checkbox"/> I understand that my identity (and that of my organisation) will remain confidential and, where appropriate, a pseudonym will be used in referring to information that I provide |
| <input type="checkbox"/> I understand that the information that I provide may be included in an institutional report | <input type="checkbox"/> I consent to take part in this research. |

Name
First Name Middle Name Last Name

Signature Date
MM DD YY

 **Consent Form: Evaluator**
Study: A systems-based approach to strategic planning, monitoring and evaluation in ECD: Designing and testing the SAMEL Toolkit
Researcher: Jessica Horler



You have been invited to participate in a doctoral research project conducted by Jessica Horler from the University of Cape Town. This project will design and test a systems-based toolkit for the development of a monitoring, evaluation and learning (MEL) framework for use by nonprofit organisations in the Early Childhood Development sector.

You have been contracted to implement a pilot testing of this toolkit. The objective of this process is to run through the toolkit via five workshops with an ECD NPO. During these workshops, you will use the toolkit and resources provided to develop a MEL framework for the NPO.

Your participation in this pilot involves: reading the toolkit, preparing for the workshops, conducting the workshops, and producing finalised outputs following the workshops.

It also involves providing written feedback in the toolkit, and in-person or telephonic feedback regarding your experience of the workshop and implementing the toolkit, following each workshop.

Please take note of the following:

- While your participation in this pilot is contracted and paid, the provision of feedback and data is voluntary.
- You have the right to refuse debrief sessions, phone calls etc to provide feedback.
- The workshops that you conduct will be audio-recorded.
- Debrief phone calls and informal conversations may be audio-recorded.
- Photographs will be taken of you during the workshop.
- You will **not** be identified by name in the dissertation report, nor any associated reports.
- Information that you provide as part of this research will be subject to standard data use policies which protect anonymity of individuals and institutions.

Please indicate your consent for this research by responding to the conditions below:

- | | |
|---|--|
| <input type="checkbox"/> I understand the purpose of the research. | <input type="checkbox"/> I understand that my feedback will be recorded so that the researcher may accurately portray my responses. |
| <input type="checkbox"/> I understand that I can withdraw my feedback at any time. | <input type="checkbox"/> I understand that my identity (and that of my organisation) will remain confidential and, where appropriate, a pseudonym will be used in referring to information that I provide. |
| <input type="checkbox"/> I understand that the information that I provide may be included in an institutional report. | <input type="checkbox"/> I consent to take part in this research. |
| | <input type="checkbox"/> I consent to de-identified photographs of me being shared in an institutional report (any identifying features will be blurred or obscured). |

Name

First Name

Middle Name

Last Name

Signature

Date

MM

DD

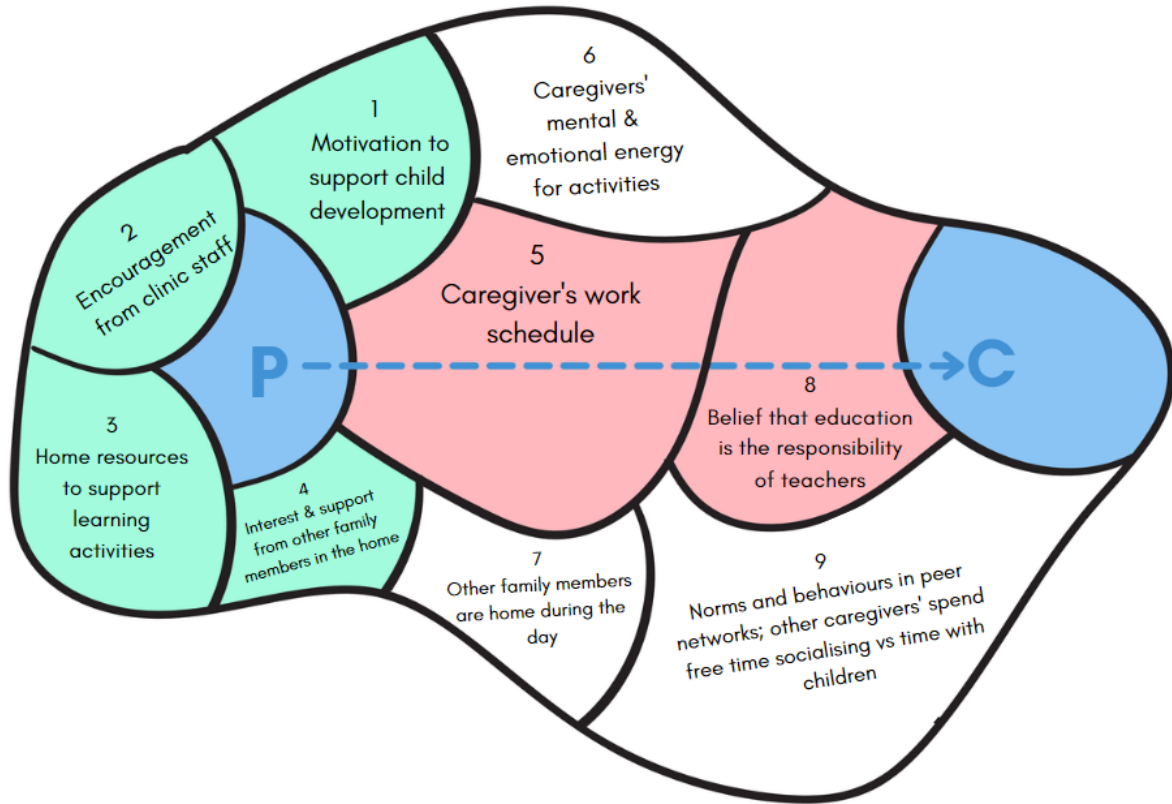
YY

Appendix Q: Life Space Example

An ECD parenting programme intends to create awareness of the importance of the home learning environment for child development (mechanism), in order to encourage caregivers to conduct early learning activities in the home (outcome).

As per the life space model below (), participants (caregivers) may be supported to achieve this mechanism and outcome with the help of: (1) their motivation to support their child's development; (2) encouragement from clinic staff to engage their young children in early learning stimulation; (3) resources in the home that can be used in early learning activities; and (4) interest and support from other family members in the home. A potentially obstructive condition is (5) the caregiver's work schedule – do they actually have the time to engage in these activities with their child? They might have the time, but they may not have (6) the mental or emotional energy after work. This condition, in turn, might interact with (1) their motivation to support their child's development – the more motivated they are, the more easily they might be able to find the energy. Caregivers' work schedule may also interact with another contextual condition – (7) the availability of other family members during the day to engage in early learning activities with their child – i.e., if the caregiver is not available – is someone else? This condition, in turn, would interact with (4) the interest and support of the other family members with regard to engaging in these activities. A second key obstructive factor may be that caregivers (8) believe that education is the sole responsibility of teachers. This is likely to interact with (9) the norms and behaviours of their peer networks – do other caregivers engage in early learning activities with their children?

Life Space Example of a Parenting Programme.



Appendix R: Certificate of Editing



This serves to certify that I duly edited:

How feasible is a systems-based approach to strategic planning, monitoring and evaluation in non-profits? Designing and testing the SAMEL Toolkit

by

Jessica Horler

I am an accredited editor with the University of Johannesburg, University of Stellenbosch Business School, NWU, UP, UCT, and GIBS, and my clients include the United Nations Global Compact, Absa, FNB, Takealot, and various other universities and organisations in South Africa and Namibia.

Please note that all editing is done in *Track Changes*, and I therefore have no control over what is accepted or rejected by the author. Furthermore, I have no control over text added at a later stage.

Should there be any queries, please contact me on the number provided above.

Teresa Kapp