Systems Thinking in International Education and Development

NORRAG SERIES ON INTERNATIONAL EDUCATION AND DEVELOPMENT

Series Editor: Gita Steiner-Khamsi, *Network for International Policies and Cooperation in Education and Training (NORRAG)*

The NORRAG Series on International Education and Development reflects the mission of the global Network for International Policies and Cooperation in Education and Training (NORRAG) to produce, disseminate and broker critical knowledge on topical issues that emerge in education and development. Through its programmes, knowledge production and dissemination, NORRAG contributes to enhancing the conditions for participatory, informed and evidence-based policy decisions that improve equity and quality of education. This series aims to act as a knowledge broker at the interface between research, analysis, policy and practice within the comparative, development and international education community.

Titles in the series include:

The State, Business and Education
Public–Private Partnerships Revisited
Edited by Gita Steiner-Khamsi and Alexandra Draxler

Philanthropy in Education
Diverse Perspectives and Global Trends
Edited by Natasha Y. Ridge and Arushi Terway

Realizing the Abidjan Principles on the Right to Education
Human Rights, Public Education, and the Role of Private Actors in Education
Edited by Frank Adamson, Sylvain Aubry, Mireille de Koning and Delphine Dorsi

Systems Thinking in International Education and Development
Unlocking Learning for All?
Edited by Moira V. Faul and Laura Savage

Systems Thinking in International Education and Development

Unlocking Learning for All?

Edited by

Moira V. Faul

Director of Research and Executive Director, NORRAG, Geneva Graduate Institute of International and Development Studies, Switzerland

Laura Savage

Associate Member, Research for Equitable Access and Learning (REAL) Centre, University of Cambridge, UK

NORRAG SERIES ON INTERNATIONAL EDUCATION AND DEVELOPMENT

Cover image: Designed by Anouk Pasquier (source file from starline / Freepik)

Contents

Figures

Tables

Boxes

Contributors

Donvan Amenya is a Research and Insights officer with the Education Development Trust. He is an Education Specialist with expertise in research and programming within the international development and humanitarian contexts. He is currently supporting the design, implementation and delivery of large-scale, multi-country educational research programmes, utilising evidence-based insights to inform thought leadership and support decision making on existing programmes.

Ailsa Buckley is the Global Guardian for Inclusive Systems Development at Swisscontact and has been designing and implementing enterprise projects in Africa for over 30 years. She has extensive experience in systems development approaches and performance measurement for systemic markets work, as well as capacity strengthening of individuals and organisations in systems thinking and practice.

Clio Dintilhac is a Senior Programme Officer on the Global Education team at the Gates Foundation. Prior to joining the foundation in 2016, Clio was a Senior Associate at the Boston Consulting Group in France and Indonesia and worked for the Africa Governance Initiative in Ethiopia. Clio has a Master's in Public Administration in International Development from the Harvard Kennedy School and a Master's in Quantitative Economics and Finance from HEC Paris.

D. Brent Edwards Jr is Associate Professor and Graduate Chair in the Department of Educational Foundations at the University of Hawaii. He has published widely on the global governance of education, education policy and political economy, focusing on middle- and low-income countries. He is the principal investigator for a three-year $913,000 project funded by the Dubai Cares Foundation entitled, 'Crisis Management for Disaster Risk Reduction in Education Systems: Learning from the Elaboration and Integration of Technology-Focused Strategies in El Salvador, Honduras, and Colombia'.

Moira V. Faul is Director of Research and Executive Director of NORRAG at the Geneva Graduate Institute of International and Development Studies, Switzerland. Her research advances systems and decolonial understandings of education, development and sustainability, highlighting questions of power, inequality and participation.

Katie Godwin is the Research Lead for the Education Commission's Education Workforce Initiative. She is from the United States and is currently based in Madrid, Spain, working as a consultant in global education research and policy. She led the development of EWI's flagship report, 'Transforming the Education Workforce', and supports EWI's in-country work to develop education workforce reform options with Ghana, Sierra Leone and Vietnam.

Abe Grindle is a Director on Co-Impact's programmes team, where he helps to oversee the sourcing, vetting, structuring and support of philanthropic systems change grants. Prior to joining Co-Impact, Abe spent almost a decade with the Bridgespan Group, where he directed teams advising leading philanthropists and NGOs operating in Africa, the Caribbean, India and North America. Abe received an MSc in Technology and Policy from MIT, where he also completed an MSc in Aeronautics and Astronautics.

Charlotte Jones is Global Head R&D at Education Development Trust, responsible for the development of evidence-informed solutions and new innovations across the charity. With a background in change management and organisational effectiveness, she is passionate about giving policymakers the practical tools to deliver on their promises.

Michelle Kaffenberger is Deputy Director of Research with the RISE Programme at the University of Oxford where she leads research on education systems, learning outcomes and ways to address the learning crisis. Previously, she served as Senior Advisor and Research Director with a global partnership at the World Bank. Michelle has designed, led and advised on dozens of research initiatives with large philanthropic organisations, multilateral donors and government ministries in low- and middle-income countries.

Mike Klassen is a Senior Associate at the Springfield Centre and a PhD Candidate in Higher Education at the University of Toronto.

Keith M. Lewin (PhD, MSc, BSc, FAcSS, C.Phys**)** is the Emeritus Professor of International Development and Education at the University of Sussex. He has published over 200 journal articles, monographs and books and was the founding Director of the international Master's programme. He has worked across Asia and Africa for over 45 years for bilateral and multilateral agencies including DFID, AusAid, the World Bank, the Africa Development Bank, UNICEF and UNESCO. He was director of the DFID funded Research Centre on Educational Access and Equity (www.CREATE-rpc) and was Chair of the Trustees of the UK Forum for International Education and Training (UKFIET) until 2020.

Varja Lipovsek is the Director of Learning, Measurement and Evaluation at Co-Impact, where she is responsible for shaping and implementing Co-Impact's approach to measurement and learning, supporting measurement efforts among partner organisations, and enabling Co-Impact to contribute to overall knowledge on systems change. Varja has worked at MIT's Governance Lab, at Twaweza East Africa, at Population Services International and at the World Health Organization. She holds a PhD from Tulane University.

Raphaelle Martinez joined Global Partnership for Education in 2013. She currently leads the Education Policy and Learning team that supports the use of evidence and good practices for education system strengthening. Previously, Raphaelle oversaw the GPE team on data, planning and finance. She has been working in the field of education for more than 15 years, of which ten were as an education planning specialist. Raphaelle holds two Master's degrees from the Sorbonne University in political science and in education planning.

Ashleigh Morrell is the Programmes and Partnerships Director for TaRL Africa, previously the Associate Director of Policy at J-PAL Africa. It was during her time at J-PAL Africa that she first started supporting TaRL programming in 2015 in Zambia, after initial discussions with the Ministry of Education revealed their interest in an evidence-based approach to help children catch up on foundational skills. Subsequently, she has supported TaRL programming in over ten countries in Africa.

Jean-Pierre Mugiraneza is the Leadership for Learning lead at Education Development Trust Rwanda. He is an educationist and economist practitioner, and he spearheads the leadership for learning system in Rwanda in the Building Learning Foundations programme, to ensure

school and system leaders have the right skills and competencies to improve inclusive learning outcome in schools. He leads the establishment and the organic growth of networks of school leaders in Rwanda.

Juan David Parra is Assistant Professor of the Institute of Education Studies from Universidad del Norte (Colombia). He has participated in several evaluation studies, being one of the precursors of realist evaluation (in education) in Latin America. In 2022 he joined an international consortium led by the University of Notre Dame as a technical advisor of USAID's Supporting Holistic and Actionable Research in Education (SHARE) initiative to advance education learning priorities in low- and middle-income countries.

Devyani Pershad leads international collaborations at Pratham and is Technical Advisor at TaRL Africa. She also leads communications at Pratham and has been with the organization since 2013.

Laura Poswell is the Executive Director for the Abdul Latif Jameel Poverty Action Lab (J-PAL) Africa, based in SALDRU at the University of Cape Town. She works closely with policymakers, civil society partners, researchers and donors to promote collaborations to increase the use of evidence in decision making and the scale-up of impactful social programmes in sub-Saharan Africa.

Sandra Rothboeck was the Global Head Thematic Scope and Skills Development at Swisscontact until 2021. For the last 25 years she has worked in academia, as a senior skills and employability specialist in the ILO South Asia and South East Asia, and as a social entrepreneur in India. Today, she is a freelance consultant based in Asia. Sandra holds an MA in labour economics and a PhD on adult learning from University of Zurich.

John Rutayisire is a former Director General of the Rwanda Education Board and was a member of the Rwandan Ministry of Education senior management team for over 15 years. He supported the establishment of the Leadership for Learning model in Rwanda and is currently working as a World Bank Consultant supporting the Inclusive Education Initiative.

Laura Savage is respected within the global education community for her thinking on education systems, foundational learning and the politics of aid. She is currently Executive Director of the International Education Funders Group, a network of 100 philanthropic actors working to

improve global education. Previously, Laura was Senior Education Adviser at the UK Foreign, Commonwealth and Development Office. Laura's approach is captured in her PhD on national ownership in the Malawian education sector.

Marla Spivack is the Research Manager of the RISE Programme and a Research Fellow on the Building State Capability programme at the Center for International Development at Harvard University. She has worked on social protection, rural development and micro-credit programmes with government agencies in Southern Africa and South Asia and holds a Master's in Public Administration in International Development (MPA/ID) from the Harvard Kennedy School.

Gita Steiner-Khamsi, Dr phil., is Professor of Comparative and International Education at Teachers College, Columbia University, New York (fall semesters) and was NORRAG's Academic Director from 2017 to 2022. In addition, she is a past President of the Comparative and International Education Society. Her areas of research include comparative policy studies in education, policy borrowing and lending, and global governance in education.

Nico Vromant is the Country Programmes Manager for VVOB – Education for Development in Zambia. Since August 2017, he has worked with the Zambian Ministry of Education towards the growth of TaRL in Zambia.

Tjip Walker is a Senior Policy Advisor in the Bureau for Policy, Planning and Learning (PPL) at the US Agency for International Development (USAID). In this role he is leading Agency efforts to implement the Local Systems Framework, USAID's commitment to sustained development that puts local systems at the centre of its work. In support of that effort, Dr Walker works across USAID to identify promising practices, develop tools and spread system thinking.

Elena Walls is Senior Monitoring, Evaluation and Learning Advisor with the Center for Education at USAID, Washington. In this capacity, she leads multiple strategic initiatives and provides technical assistance to USAID missions. Before joining USAID, Elena served as Director of Monitoring, Evaluation and Research team at Education Development Center. She holds a Doctorate degree in Sociology from the University of Maryland, College Park, and an MA in Society and Politics from Lancaster University, UK.

Jason Weaver is a Senior Economist with the Education Global Practice of the World Bank. His operational and technical work has covered many countries across Africa, Europe and Central Asia. Jason is currently leading the design and launch of the World Bank's Accelerator programme to tackle Learning Poverty and the forthcoming regional education strategy for the Bank's engagement in Western and Central Africa. He received his degrees from The Ohio State University and Harvard University.

Foreword

Gita Steiner-Khamsi

This is the fourth volume in the *NORRAG Series on International Education and Development.* The series aims to serve as a knowledge broker at the interface between research, analysis, policy and practice within the comparative, development and international education community. The books in this series intend to generate an international debate on emerging trends in education and provide space for authors that represent diverse perspectives and knowledge communities.

The first volume opened new horizons on the debate about public-private partnerships in education as they diversify, expand and increasingly blur the distinction between the public and private sectors. The second volume explored trends, controversial issues and new forms of philanthropy in education. The third presents multidisciplinary approaches to explain the development of the Abidjan Principles, including their articulation of the right to education, the state obligation to provide education, and the role of private actors in education.

This fourth volume is likely to become the authoritative book on systems thinking and systems design in education. Today we are witnessing a plethora of platforms that intend to offer good advice on how to fix schools based on a universal, one-size-fits-all approach. This book convincingly shows why such an approach, often sailing under the flag of 'international standards', 'best practices' or 'what works', has failed miserably. In practice, this has often led to disruptive and chaotic reforms in which the 'innovations' are merely added to existing practices (often only for the duration of external funding) without replacing dysfunctional old structures and practices. The authors of the book offer a more nuanced and complex approach that takes into account the interplay of various actors and elements within the system. To carry out effective reforms, we need to start thinking like a system.

<div align="right">

Gita Steiner-Khamsi
Series Editor, New York and Geneva

</div>

Acknowledgements

The authors of this volume would like to thank NORRAG for this opportunity to showcase the emerging field of systems thinking in international education. The year-long journey towards the compilation of this volume has revealed the broad and deep range of efforts to apply systems thinking theory to education practice and the lessons, in turn, that education practice can yield for systems thinking theory. As a group of editors and authors, we have expanded our own knowledge of systems thinking in international education and found a wider network of peers with whom to debate and further the field. We would also like to thank Georgia Thorne, Daniela Andonova and Anouk Pasquier for their patient and careful coordination and copy-editing of the volume.

1. Introduction to *Systems Thinking in International Education and Development*

Moira V. Faul and Laura Savage

Too many education systems around the world are failing. Large investments from low- and middle-income country (L&MIC) governments and high-income country donors have increased the number of children attending school. This is an incredible achievement. Despite decades of investment and initiatives, the children now in school were the easiest to reach; now that they are in school, it is clear that too many are not learning enough, much less learning to love learning (Lewin, 2009; Little & Lewin, 2011; Kaffenberger & Pritchett, 2021). In 2019, 53% of ten-year-olds worldwide were unable to read and understand a short age-appropriate text (World Bank, 2019). Student learning trajectories in some countries have been declining over the past decades, despite significant financial investment and reform efforts (Kaffenberger et al., 2021; Le Nestour et al., 2021). One survey conducted in Uganda in 2013 showed that only one in five teachers understood the content of the curriculum they taught (Wane & Martin, 2013). The existing learning and equity crises have been further exacerbated by the COVID-19 global pandemic, causing the most widespread school closures in history. Rhetoric within global education discourse grows ever more frantic, with major efforts in 2022 to build back better, reimagine the futures of education and avoid losing a generation (Steer, 2020; UNESCO, n.d.; Ahlgren et al., 2022). These statistics and reports underline the urgency and scale of what is known in international education circles as the 'learning crisis'. Current so-called common sense or expert solutions have not even led to improved basic literacy rates, much less many of the other education and life outcomes promised in Sustainable Development Goal (SDG) 4: to 'ensure inclusive and equitable quality education and promote lifelong learning opportunities for all' (UNGA, 2015).

The complex nature of this crisis does not seem fully appreciated. In this volume, we explore the twin global learning and equity crises as a 'wicked' problem, to which 'there are no "solutions" in the sense of definitive and objective answers' (Rittel & Webber, 1973, p. 155). In this, we declare our mental model up front. We believe that to 'ensure inclusive and equitable quality education and promote lifelong learning opportunities for all' (UNGA, 2015) fundamentally requires systems approaches. The dream is that all children, youth and adults everywhere learn and love learning, not just to get a job but because a positive journey of lifelong learning is how humans thrive. However, we are so far from that goal that we are focusing here on the fundamental change needed to begin to address this wicked problem whether at classroom or global scale: from a linear view of change to a systems view.

One of the virtues of systems thinking that we value is double loop learning, or the ability to question our assumptions and frameworks. At the same time as believing in the goal of foundational learning for all, we are concerned about the risk in the focus on and measurement of so-called basic learning outcomes: towards children becoming objects on a conveyor belt into an economic growth-only education model. We believe that education systems should enable and encourage children to enjoy our broad understanding of quality and inclusive learning for all. However, our systems thinking approach means that we recognise that every actor within an education system will have their own goals for education and that powerful incentives and interests lie within every system that shape what emerges. If we deny this complexity, unintended consequences multiply. It is far better to acknowledge the complexity of the systems in front of us and work with them.

Two questions are becoming more urgent as the 2030 target date for the Sustainable Development Goals looms. First, we have to ask why? Why are children not learning the very basics of literacy and numeracy, despite the depth of evidence on how to do this? Why do interventions resulting in significant learning gains on a small scale not translate into large-scale success (Bold et al., 2018)? Why are there so few examples of large-scale, let alone sustainable, learning gains?[1] The second question is how? How do education actors need to act differently to achieve SDG4, at a pace and scale not yet seen before? So-called scaling requires respect

---

[1]   A search for eight examples of projects that successfully improved learning at scale struggled to identify candidates (Stern et al., 2021).

for the local system into which the initiative intervenes (Faul, 2016a). Whatever the globally produced evidence base, what is required is locally relevant evidence and systemic rewards for using it (Steiner-Khamsi et al., 2022).

Thus, we argue that answering the 'how' question also addresses the 'why': taking a systems approach towards educational change offers promising responses to these two vital questions. Systems thinking provides an approach that is now established within the international development literature. The exponents of systems thinking hold prominent roles in the efforts to tackle the climate crisis (Clinton Foundation, 2019). Those working on energy and agriculture in low-income countries speak with fluency about food systems (IFPRI, n.d.). An established community of self-identifying systems thinking researchers and practitioners seeking to hasten the implementation of systems thinking into agricultural practice (Nature Food, 2020) has applied lessons from systems literature into the recent United Nations Food Systems Summit (Borman et al., 2022; Brouwer et al., 2020; Dekeyser et al., 2020; Ericksen, 2008; Ingram, 2011; Posthumus et al., 2018). The debate at the biannual health systems conference draws together researchers, policymakers and practitioners to build awareness of how health systems work (Health Systems Global, n.d.), allowing lessons from systems research in health to start being applied to education (Hanson, 2015). In sector-agnostic discourse on effective development, works by Ben Ramalingan (2013) or Jean Boulton et al. (2015) on complexity and Duncan Green (2016) on how change happens are discussed with familiarity and ease at dinner parties among aid workers. Furthermore, in interviews with Ministry officials in low-income countries, Faul (2018) found enthusiasm for a systems approach and a sophisticated understanding of the implementation realities that make this necessary.

The word system is becoming more widespread in the global education discourse. We hear many education stakeholders using the word system when they used to say sector. Rather than talking about education sector reform, global education actors increasingly talk about 'education system change' and anchor their strategies around supporting education 'system transformation' (e.g., GPE, 2020). The 2018 World Development Report, the first ever of this World Bank flagship series to focus on education, set out a framework for 'how to make systems work for learning' (World Bank, 2018). Terms such as 'systems approach', working 'at the system level' and claims of supporting 'system change' are commonplace. Nevertheless, not all of these translate into the application of systems

thinking; a discursive shift does not mean that there has been a shift in behaviour or practice.

The meaning of and implications associated with this change of vocabulary from sector to systems are not yet well characterised, much less implemented. Thus, there is a real danger that the change in vocabulary away from sector and to 'systems' may do no more than mask the perpetuation of practices that have been repeatedly shown to be inadequate for decades. In this book, we highlight theory and practice from representatives of a small but growing community who are not just using the word 'system', but who are turning to the core tenets, theories and tools of systems thinking to drive progress towards unlocking equitable and inclusive learning for all. We provide some shining examples of efforts to apply systems thinking in global education – if not yet much research on whether and how this has changed mindsets or decisions more widely. The chapters in this volume highlight these examples, presenting diverse systems approaches currently being theorised and applied to education in middle- and low-income countries. Through these examples, this volume showcases the range of innovation in the emerging field of education systems in research and policy; collates new frameworks and tools for understanding education systems; and shows how this matters for policy and practice.

This introduction has two main parts. First, we provide a brief tour through the literature that has informed global education policy and practice in recent years. We demonstrate how this has led some individuals to seek answers in systems thinking. Second, we introduce the core concepts of systems thinking and translate this vast and sometimes dense literature into implications for international education. We use the term 'international education' as shorthand for the efforts that actors around the world are taking to investigate and achieve SDG4.

WHAT DO WE KNOW ABOUT HOW TO IMPROVE LEARNING FOR ALL?

Given the urgency and magnitude of the international education challenge, a frequent refrain heard over the past decade from policymakers around the world is: what works to improve learning and equity? More funding has been made available for research into improving learning; some researchers have estimated that this is the second fastest-growing field in international development (Evans & Mendez Agosta, 2021). The Global Education Evidence Advisory Panel in 2020 produced a report

called Smart Buys in Education (GEEAP, 2020), which summarises the strength of evidence and cost of implementing a list of interventions, giving the positive message that so-called 'great buy' interventions can be transplanted and implemented in different places for the same results to allegedly fix the learning crisis. This is a much more compelling narrative to policymakers and funders than that of a few years previously, which was far more cautious and heard as 'we do not know enough about how to improve learning'. Whatever the 'what works' literature may contribute, it does not answer the 'why' and 'how' questions raised above: why are education outcomes so slow to improve particularly among the most marginalised, and how can we change this?

Nor does this literature address the evidence that the wealth of global expert technical solutions produced in the past decades has not produced the sought-after improvements. Permanent doubling of teacher pay in Indonesia has had no effect on student learning outcomes (de Ree et al., 2015). Inputs, such as textbooks, may be procured and delivered effectively, but this does not mean that student learning will rise (Savage, as cited in Gibbs et al., 2021). A pilot project to improve teacher incentives and performance through issuing fixed-term contracts led by a non-governmental organisation in Kenya improved student learning outcomes significantly (Bold et al., 2018). However, there was zero impact on student learning when rolled out nationwide because of lack of consideration of system factors, such as opposition from the teachers' union and government bureaucracy, which then led to implementation delays and issues with interpreting contractual terms. Was it that implementation fidelity waned when the contract teacher reform was rolled out nationwide? If implementation fidelity could be fixed, could the pilot-to-scale equation work? A large-scale experimental evaluation in India suggests not (Muralidharan & Singh, 2020). Here, the researchers tracked an ambitious reform effort, including several 'best practices', such as school improvement plans and school ratings. Implementation was completed as planned, but the reform had no impact on school management or student outcomes. Regardless, because it was implemented as planned, it was considered successful and scaled up to 600,000 schools nationally.

Why do such so-called common sense expert solutions not result in students learning? Those from the wider school of systems thinking might argue it is not surprising that the perfect implementation of the wrong intervention would result in zero or a negative impact. Linear approaches that 'plan–do–review according to project objectives' (and

not system effects or outcomes) are doomed to fail when dealing with nonlinear complexity in education systems. To achieve an emergent outcome, for example student learning in a complex system such as in education, there can be no quick and replicable solutions at scale.

Scouring the literature for education reforms that have produced learning 'at scale' finds few examples, and a general conclusion that far too little data on learning outcomes or on the proxies of education system performance is collected. A 'test–learn–adapt' approach to implementation alongside strong political will, as well as technical support, has been critical to the success of a structured pedagogy reform in Kenya (Piper et al., 2018). In common with the Kenyan study, successful reforms in Mexico, Tanzania, India and Brazil (Crouch, 2020; Stern et al., 2021) also involved several components, including providing students with books, face-to-face teacher training, coaching and support focused on practising new skills and using student assessment to target teaching content and pedagogical strategies to real, not assumed, student capability.

So does 'system change' mean doing lots of things at once and at scale? Many examples of similar interventions did not work, so what made the above examples special? Data on the question of why these projects worked are limited. One promising angle is pursued in a small corner of the global education literature focusing on the politics of education reform. Political science is not systems thinking, but it is a field that asks – and begins to answer – useful questions about the political context into which any education reform intervenes. Some of the most promising answers to the why and how questions have come out of this field. Three works in particular start to answer some of the why questions that the experiments referred to above cannot. One book specifically asks the following question: 'Why is learning achievement in South Africa so low?' This book develops a framework that explains how politics, institutions and governance interact (Levy et al., 2018). The authors suggest that any analysis of teacher training or other education interventions needs to consider how these efforts interact with the political context of the education system, in this case prevailing national and local governance systems. These governance systems deeply affect technical implementation and whether the actors are motivated to demand learning outcomes. Nevertheless, demands for implementation fidelity to globally designed reforms recur, far more frequently than demands for locally relevant and context-sensitive reforms. Second, comparative analysis of education reform in Bangladesh, Cambodia, Ghana, Rwanda and South Africa suggests that success is found among localised solutions at the school level,

often against national-level political settlements, arguing that support should build coalitions for change from the bottom (Hickey & Hossain, 2019). Third, a fascinating ethnography of an education reform in Delhi shows that reform efforts can only succeed if the 'frontline of the system' understands and supports them (Aiyar et al., 2021).

These works have contributed a great deal to understanding why education investments that ignore education systems' contexts do not move the needle on learning and equity. A systems thinking lens would underline similar conclusions regarding the importance of locally led, non-standardised and context-responsive education to enable all children to learn (and to love learning). How to understand and work with a system and its context lies at the heart of systems approaches for education.

There is a final question in the literature that focuses on the extent to which outsiders to a system can change the system: are the claims of global actors to 'drive system change' even practicable? The wider literature on aid effectiveness supports the need for country ownership or the co-creation of interventions (Faul, 2014; Savage, 2013). There are some interesting theories emerging in the international development literature that challenge the effectiveness of tight control by donors (Honig, 2018) and the need for building coalitions for change (Hickey & Hossain, 2019) or coherence for learning (Pritchett, 2015). Three conclusions are drawn from the education literature. First, education systems are expected to change too rapidly and too often by both international donors and national stakeholders, particularly those caught up in electoral cycles (Hargreaves & Shirley, 2009). Second, there is over-reliance on the wrong drivers for system reform, including on external accountability alone to deliver results (Fullan, 2010, 2011). Third, the global education aid architecture is broken (Burnett, 2019; Beeharry, 2021). Systems thinking offers the concepts and tools to understand – and act to improve – relationships between actors within education systems and interactions between different systems (e.g., the global education aid system and a national/local education system).

This section has provided a journey through some of the trends and themes of global education literature over the past ten years, showing the contribution of systems approaches to improving research, practice and results in the decade to come. Education systems research is not yet an

established research field, as it is in other sectors.[2] This volume highlights that there is some really exciting thinking and theorising happening in a two-way conversation where systems thinking informs education practice and vice versa. It also aims to stimulate the system of actors and institutions that would fund and do such research and do not yet have sufficient incentive towards this as a goal.

WHAT IS A SYSTEM, AND HOW CAN WE UNDERSTAND IT?

A system is a group of interconnected components with shared purpose that together achieve more than the sum of their parts. An education system can be a classroom, a network of teachers or a global architecture. Within each and across each lie other systems. It is not possible to predict or change an entire system. However, it is possible to identify the levers through which to encourage change within a system. There are different types of systems that work in different ways, and there are also a variety of tools for understanding the dynamics within a system. All of these are fundamental insights from the field of systems thinking.

Systems thinking is an umbrella term for ideas and practices in research, policy and practice that are derived from physics and engineering, life sciences and social sciences and informed by ontologies and theories as diverse as cybernetics, chaos and complexity. Broadly, systems thinking enables a more holistic consideration of the following:

- system elements: both material (teachers, schools) and intangible (beliefs, information);
- the relationships between those elements and subsystems;
- the structuring of the system and subsystems within it;
- the functions of the system (both formal/stated and informal/in practice); and
- positive and negative feedback loops and influence pathways in the system.

There are some useful renditions of sometimes dense systems thinking concepts into compelling and practical guides for international develop-

---

[2] Some important research has been done into complexity and education, however, including but not limited to Fenwick (2012); Lemke and Sabelli (2008); Mason (2008, 2009); Morrison (2008).

ment and social change (Meadows, 2008; Stroh, 2015; Rayner & Bonnici, 2021; Boulton et al., 2015; Bowman et al., 2015; USAID, 2016). Here, we summarise only the main concepts, giving illustrations relevant to the SDG4 education agenda. We introduce the core differences and terms set out in simple terms[3] in Figure 1.1.

Simple, complicated	Complex
Linear	Nonlinear
Closed	Open
Technical, mechanical	Social, living
Single loop	Iterative
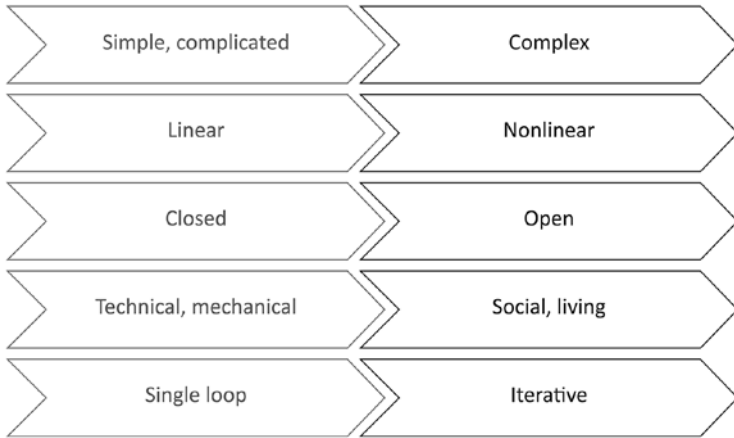
Source: Authors, based on Boulton et al. (2015), Green (2016), Ramalingam et al. (2014).

Figure 1.1   Systems thinking concepts

Complicated and Complex?

If a system is complicated, one set of actions is appropriate; a different set of actions is required if it is complex. Some systems theorists (such as Checkland, 1981; Flood & Jackson, 1991) have divided systems approaches into those that better respond to technical problems ('hard') and those that are more appropriate to human systems ('soft'). The 'hard' systems approach has the easiest analogies in technology or engineering, where there are well-defined problems, a clear goal in sight and the process of applying a systems approach leads to achieving that goal. The underlying philosophy is to 'do the thing right'. Nevertheless, if the

---

[3]   Systems thinking does not encourage such binary thinking, but this helps highlight the core tenets and virtues (to us) of systems thinking.

wrong thing is chosen, it does not matter how faithfully you implement it. In contrast, 'soft' systems approaches are better applied to messier, less-structured problems and involve a process of participation and debate. Soft systems approaches are considered successful when they result in learning or satisfaction of the participants; there is not necessarily a solution or final resolution of the problem. The underlying philosophy is to 'do the right thing' (Pan et al., 2013).

Development problems are too often understood as complicated, that is, tractable to being boiled down to issues of one or two variables, which can be dealt with by applying technical solutions (Ramalingam et al., 2014). Taking this approach in education assumes that it is a complicated problem (like plumbing or electrical wiring) rather than a complex one (a garden or weather system). The difference between these two types of systems is profound: although the elements in the first remain unchanged by their interactions with others, in the second – complex – system, the elements are changed by their interactions with other elements in the system, with unpredictable and unintended consequences. A response should always be proportionate to the nature of the problem; thus, in health, de Savigny and Adam (2009) argue that the need for systems thinking increases commensurate with the complexity of the intervention and likelihood of system-wide effects (p. 34).

As systems thinking has evolved, it has become clear that both approaches can usefully be deployed simultaneously because most systems are techno-social, that is, comprise a combination of so-called hard and soft aspects that are interlinked and frequently interdependent. For example, a soft systems approach might be applied to asking the hardest-to-reach groups with multiple vulnerabilities regarding why they cannot access or school, followed by a 'hard' systems approach to tackle issues of infrastructure (such as ramps or safe toilet facilities) and a further soft systems approach for addressing exclusionary behaviours and bullying.

Thus, systems approaches can feel threatening to experts who 'know' that they have a technical solution. Instead, systems approaches move away from best practice models to the best fit. Best fit programmes are optimally adapted to different educational, political, social and economic contexts through engagement with local actors who lead and sustain changes (Ramalingam et al., 2014). Solutions to complicated problems require expertise alone, whereas complex problems require local engagement, remaining sensitive to local differences and power relations for an intervention to have a chance of being successful and sustained.

Sustainable changes leading to improved learning outcomes cannot occur in the absence of the involvement of the individuals and groups who will implement the change (country government or local stakeholders), no matter how sophisticated the tools and processes that are designed and used. Why children are in school and not learning, why learning levels are decreasing and why expert solutions are not helping children to learn are complex problems to which a systems approach applies.

Nonlinearity and Causal Loops

Systems thinking is characterised by an understanding of the nonlinear nature of change, including feedback loops, tipping points and thresholds. In conventional (sector) approaches, change is believed to be linear, additive and proportional to the inputs and outputs. That is, there is the belief that generating more of output x necessarily leads to more outcomes and impact.

If we visualise linear change as a graph plotted against x and y axes it would show an upward or downward line: where two variables are in linear relationship, the direction of change is the same, and there is a constant rate of change of y (e.g., $x=y$, $x=2y$, $x=3y$, etc.); where one increases so does the other at the same rate, and where one decreases so does the other at the same rate. Where there is a monotonic relationship, when x increases y always increases too but not over a constant rate of change. That is, the variables follow the same direction of change (increasing or decreasing together), but at different rates, resulting in an upward or downward curve, respectively. In contrast, in non-monotonic relationships the two variables measured will not consistently either increase or decrease in relation to another. The relationship between the variables may sometimes show an increase and sometimes a decrease, and the rate of change can vary as well as the direction of change, resulting in a u-shaped or n-shaped curve or even more complex lines. A monotonic relationship can be nonlinear but x will follow the same direction of change as y. Thus, Laumann et al. (2022) argue that 'the known complexities of interconnected socioeconomic, climatological, and ecological systems, [are] characterised by non-linear, non-monotonic relationships' (p. e423).

If relationships between variables are nonlinear and non-monotonic, then only mathematical and statistical techniques that can capture this nonlinearity and non-monotonicity are appropriate. Thus, in Figure 1.2 below (Laumann et al., 2022), applying a Pearson correlation to study the

relationship between children out of school and GDP per capita in Latin American and Caribbean countries reveals a weak, non-significant linear relationship. Nevertheless, a high, significant correlation was detected when using partial distance correlations, which allow the calculation of nonlinear dependencies while discounting the effects of lurking variables. In sum, analyses of education systems must allow for and account for probable nonlinearity and non-monotonicity of relationships between education and other sustainability goals, as well as within education systems themselves.

Inside complex systems (such as education), reforms have whole system effects that do not follow linear causality but instead entail nonlinear changes and nonlinear causality, where the causes and effects may feed into each other. Identifying these causal loops reveals the patterns that hold systems in their current (mal)functioning. Systems tools allow governments and analysts to better identify these causal loops and the system levers that can be targeted to shift the system to functioning differently.[4]

Recognising nonlinearity and unpredictability as systems features (rather than bugs) is somewhat liberating. This tenet of systems thinking means that we should give ourselves a break; we cannot know what the impact of our intervention in a system might be. Pretending that we can is magical thinking and will only multiply the unintended consequences of any action we take. We can – and should – observe and learn, and then reconsider and reconfigure the actions we had planned in ways that take into account what we have observed. However, we cannot predict with certainty what outcomes will arise from which actions and inputs invested. This realisation may initially be frustrating – so much so that we deny it, and reject systems thinking. Yet time and again, as the literature review above underlines, even the most basic objective (such as enabling children to read) does not emerge from a set checklist of actions undertaken identically in every corner of the globe. We accept this in so many other aspects of life: when baking bread, we may follow a recipe and set

---

[4] It is worth noting that no lever can result in whole system change or predictably accelerate any system outcomes. Levers can be used to prod a part of a system, and lever-pullers will then need to track the feedback loops and unintended consequences that emerge from this. Of course, all sorts of levers are being pulled at the same time by different lever-pullers, many of whom will have different objectives. This lens helps manage expectations about the potential impact of a single project or change.

Latin America and the Caribbean
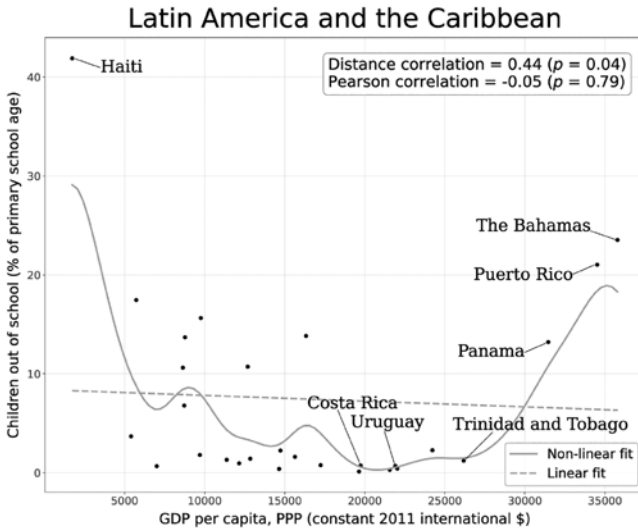
Source: Laumann et al. (2022, e424). Represented is one of the indicators of SDG 4 – children out of school (percentage of primary school age) – and one of the indicators of SDG 17 – gross domestic product (GDP) per capita, purchasing power parity (constant 2011 international $). Each point corresponds to the latest available measurement of the two indicators for a country in this grouping. The dependence between the indicators is captured more accurately through a nonlinear, non-monotonic relationship rather than a linear one, as shown by a high, significant distance correlation versus a weak, non-significant linear (Pearson) correlation. The nonlinearity and lack of monotonicity is reflected in a synergistic relationship up to about 20,000–25,000 international $ of GDP that turns detrimental with higher values of GDP (for Panama, Puerto Rico and The Bahamas).

Figure 1.2   *Example of a nonlinear relationship between two indicators for SDG 4 and SDG 17 from the World Bank (2020) data set for countries in Latin America and the Caribbean*

the oven to a particular temperature, but the resulting loaf will look and taste different depending on the protein level of the flour, the humidity of the air and the adeptness of the baker's knead. It is possible to do good for millions of learners; it is just impossible to predict what good may emerge from the myriad of acts of attempting to do good for children in different contexts with different initial conditions.

Emergence and Systems Dynamics

Systems thinking seriously considers the concept of emergence, or the 'arising of novel and coherent structures, patterns and properties during the process of self-organization in complex systems' (Goldstein, 1999, p. 49). Emergent properties are unexpected consequences that arise from the relationships between different elements of a system, resulting from the collaborative functioning of a system. Nevertheless, they do not belong to any part of that system. Take this introductory chapter as a complex system. The editors set the boundary of this system, the topic and word count, and the readers expect the writers to follow certain conventions regarding content. Words are single components that do not transmit much sense by themselves. Joined together in sentences and paragraphs, they communicate our message, and (we hope) an emergent understanding of how systems thinking can be a useful lens to under-standing and acting on ensuring quality learning for all. However, the editor and co-authors have no control over your, the readers', emergent interpretation. The emergent meaning you take away is shaped by your existing knowledge and perspective, which is likely to change upon future experiences or discussions.

Systems thinking enables a more sophisticated understanding of prob-lems and system dynamics by focusing on the interactions and feedback between components and emergent properties. The concept of dynamic complexity indicates that the behaviours – and outcomes – of a system arise from the interactions between its elements over time. The behaviour of the system and of the elements within the system are likely to change over time, and this must be accounted for. For example, studies have shown that initial score increases in high-stakes assessment are not main-tained over time, even where the same inputs are maintained (see, e.g., Wyse & Torrance, 2008).

SYSTEMS TOOLS

Conventional methods cannot capture these dynamic processes and relationships or nonlinear causality. Nor can they identify and portray the positive and negative feedback loops that generate unintended conse-quences, or the emergent properties that arise from system interactions. There is a whole suite of systems thinking tools that help with these different challenges (Table 1.1).

Table 1.1   *The value-added of systems tools*

Feature	Conventional response	Systems tools and contribution
Systemic challenges	Tendency to see development problems as ones that can be boiled down to issues of one or two variables, which can then be dealt with through applying specific *technical* solutions	*System dynamics and systems thinking* approaches enable a more sophisticated understanding of problems, by focusing on interactions and feedbacks between components and emergent properties
Behavioural challenges	Assumptions of perfect *rationality* loom large in development. It is believed that by changing knowledge, one can change behaviours	*Adaptive, agent-based models and behavioural techniques* focus on simulating 'experiments' in trial-and-error behaviours, to strengthen outcomes
Relational challenges	Much of aid analysis *underestimates relationships* and networks between actors, and focuses instead on individualised actors and entities	*Network analysis* allows the mapping of the relationships between system actors or elements and the analysis of how the relationship structure affects behaviours
Dynamic challenges	Change is seen as *linear*, additive and proportional to inputs and outputs, such that generating more of output X will lead to more outcomes and impacts	*Dynamic analytical models* help further an understanding of the nonlinear nature of change, including feedback loops, tipping points and thresholds

Source: Adapted from Ramalingam et al. (2014, pp. 11–12); Faul, 2016b.

Some of the better-known tools are causal loop diagrams and the iceberg model. Such tools enable users and participants to capture context-specific, nonlinear, emergent and dynamic processes and relationships. In other words, they help us understand the complexity well enough that it ceases to be mind-boggling and frustrating. Instead, it becomes something we can make sense of and use to design and implement actions or programmes that might work for that system.

For example, in the gap between the widespread contemporary use of systems language and the low incidence of systems thinking practices, we can find competing mental models. Saying the word system instead of sector is easy, but it cannot contribute to achieving SDG4 if practices

do not change; treating education as a system can have, and has had, profound implications for the professional behaviours of decision makers, practitioners and researchers. The insight that different mental models are at the core of this education systems conversation is critical in ensuring that participants do not talk past each other and become frustrated: they may be seeing the same events and patterns but interpreting them differently. With a systems approach, no silver bullets are possible; the very word solution is discouraged. Linear explanations of change do not make sense; thus, the log frames attached to sector plans or donor projects are no more than fairy tales. Systems thinking in international education remains contentious because it refuses to tout a single, one-size-fits-all solution; that can never be adequate to the complexity of the learning and equity crises that learners face. Instead, it recognises the reality of the complex messiness of the education system(s) it is addressing, encouraging us to take an approach focused around short feedback loops of action–feedback–change, to break out of sector-based siloes and to constantly strive to understand other system actors' perspectives, as we will see in case studies elaborated throughout this book.

Researchers, practitioners and actors within the global education architecture have developed various frameworks in recent years that translate some of these concepts into actionable formats. The UK Economic and Social Research Council ran three research calls around (1) system elements, (2) system dynamics and (3) system context (Magrath et al., 2019). The World Development Report of 2018 places learning as the emergent property at the centre of concentric circles depicting the interaction between learners, teachers, school inputs and school management, with a whole range of influencing actors around them (World Bank, 2018). The Education Development Trust identifies six accelerators to achieve at-scale educational improvement, including data, evidence-informed policy, vision and leadership and school leadership effectiveness (Ndaruhutse et al., 2019). We have engaged with at least three local education organisations that have used systems thinking tools to understand the system they work within and devise strategy based on those emergent insights (STIR Education, Central Square Foundation and Educate!) At the same time, Faculties of Education are training students in applied systems thinking in global education courses (American University, 2020; Penn GSE, n.d.; and the Education University of Hong Kong, among others). Systems thinking terminology is catching on across the mainstream international education community.

WHAT THIS BOOK CONTRIBUTES

The authors contributed chapters to this volume that address a wide range of systems approaches in education, offering case studies in countries as diverse as Bangladesh, Brazil, Cambodia, Colombia, Democratic Republic of Congo, Ghana, India, Indonesia, Kenya, Rwanda, South Africa and Zambia as well as at the global level, that raise questions of ontology, methodology and practice. Their intention is to share new insights into how education systems work (and the system boundaries are drawn differently in each chapter) while suggesting methodological and theoretical approaches towards understanding them.

Following this introduction, the chapters in Part I draw insights from findings to improve the understanding of how education systems work using systems approaches. In Chapter 2, Varja Lipovsek, Laura Poswell, Ashleigh Morrell, Devyani Pershad, Nico Vromant and Abe Grindle reflect on their experiences of implementing and researching teaching at the right level (TaRL) projects in Zambia. The practice of TaRL follows systems thinking principles of iterative tight feedback loops between action, effect and reaction to the data on those effects, in this case by grouping students according to their learning needs, teaching at the level of the group, assessing learning, then regrouping students according to their learning needs, and so on. This systems practice maintains a tight feedback loop between meeting students where they are and where they may move to. In Chapter 3, Charlotte Jones, John Rutayisire, Donvan Amenya, Jean-Pierre Mugiraneza and Katie Godwin explore how peers within education systems can be encouraged to interact and create networks of collaborative professionalism and leadership from the middle. Drawing on case studies in Kenya, Rwanda and India, the authors show that system change can arise from those processes that encourage trust, transparency and challenge among peers, alongside accountability and authorisation from mid-level bureaucrats. Chapter 4 reports on a funder roundtable convened at the 2021 Conference of the Comparative and International Education Society (CIES). In contrast to the narrative that systems approaches are anathema to funders and policymakers (although not to NGO staff and teachers), representatives from major state and philanthropic funders discuss their perspectives and priorities for furthering systems practice, beyond rhetoric.

The chapters in Part II apply insights from theory to systems thinking and practice. In Chapter 5, Gita Steiner-Khamsi writes a forceful case

for applying Luhmann's sociological systems theory to education to better explain why, when and how education systems respond to (quasi) external pressure (such as globalisation) to justify domestic reforms. She explains the concept of travelling reforms to explore how to identify the particular moments in which systems are ready to consider observing – or even internalising – education reform experiences from elsewhere. She challenges us to think about what a system is, how to work in a state of constant flux and that there is far more nuance to the pilot-to-scale debate than is usually recognised. In Chapter 6, Keith Lewin draws upon another UK government-funded research programme, CREATE (n.d.), that ran across eight countries to present a new way of thinking about education systems. He identifies the zones of exclusion for a student, presenting their journey through schooling as dynamic with several transition moments. Seeking to understand a system from the perspective of an individual student going through it is novel, and this approach prompts thinking on the multiplicity of pathways possible through an education system (against conventional narratives of a linear journey, where we need to simply push children back on track) and on the diversity of education systems outcomes (beyond literacy) pursued. In Chapter 7, Michelle Kaffenberger and Marla Spivack introduce a theoretical framework to understand how the elements, accountability relationships and feedback loops within education systems might interact to enable or limit student learning. This is based on a large UK-funded research programme with long-term research teams in seven low- or middle-income countries, each of which has tracked major education reforms using a range of methodologies (RISE, n.d.). In this chapter, Kaffenberger and Spivack explain how this framework allows for analysis of accountability relationships across an education system and, through this, an understanding of whether a system is 'coherent' for learning.

In Part III, the chapters report on the application of systems approaches or tools to practice. In Chapter 8, Mike Klassen, Sandra Rothboeck and Ailsa Buckley apply an inclusive systems development lens to technical and vocational training projects, favourably comparing this to the market systems approach that tends to be more commonly applied. They argue that vocational training systems are often defined too narrowly when it comes to incorporating changes in other parts of the system that critically affect the success or failure of vocational reforms. A deep dive into a case study from Cambodia reveals the potential benefits of the inclusive systems lens, where system failures are understood first from the perspective of those who are excluded or failed by that system, which, in

turn, aids understanding and addressing system dysfunction. In Chapter 9, Juan David Parra and Brent Edwards explore how realist evaluation can use systems thinking. They apply systems thinking to the rollout of a large-scale education reform in Colombia, finding that the method allows researchers to explore the aspects of systems thinking (such as system boundary setting and relationships between nodes) while using a language that policy makers can readily interpret. In Chapter 10, Elena Walls and Laura Savage describe their experience in designing and supporting an education system diagnostic in Rwanda and the Democratic Republic of Congo, where participatory workshops using systems thinking tools led to new insights among the participants regarding each other's perspectives and incentives, and ultimately informed USAID programme design.

MOVING FORWARD WITH RESEARCH AND PRACTICE

This edited volume contributes to the current literature in at least three ways. First, the volume curates an interdisciplinary collection of diverse systems approaches that are currently employed in education research, policy, practice and planning. The volume brings different systems approaches into dialogue with each other, thus contributing a compilation of different dimensions of what is – and is not – a systems approach in practice, including complexity, relations, nonlinearity, emergence, feedback loops, tipping points, multiplicities of actors, perspectives and motivations. To the best of our knowledge, this is the first analytic anthology that allows researchers to examine systems approaches in use in international education from multiple angles, and that addresses the contribution of systems approaches to achieving SDG4 by targeting learning and teaching directly or through research, policy and planning. Second, the volume presents a range of interdisciplinary theories and methods for researching and implementing systems approaches that demystify systems approaches and allow readers in research, policy and practice to select the appropriate strategies and methods for achieving their desired outcomes. Finally, this volume presents new real-world data and case studies on systems approaches in education, offering empirical studies and reflections on how actors have navigated, researched and tackled systems change.

Demand is increasing for an extended treatment of systems approaches in education, as well as for theoretical and methodological tools with

which to analyse them. This book is – to the best of our knowledge – the first to consider the nascent field of global education systems research. Unlike the climate, agriculture and health fields, education systems research has no journals, conferences or formal communities focused on the topic. This makes this book an important contribution to the literature, both on international education and the wider systems thinking in development literature, while also serving as a rallying point for this field; a field that has already caught the attention of funders, researchers and practitioners. This edited volume brings together theoretical, methodological and contemporary empirical work exploring the use of systems approaches to understand and influence education systems and improve learning outcomes. We hope this provokes further discussion on the utility of systems approaches for analysing and changing education policy, practice and outcomes; the limitations of such approaches; and the likely next steps for collaborative research and practice.

Our aim is to advance the debate on how systems approaches can help in understanding contemporary global education crises and showing some useful ways forward. We would like this book to be thought-provoking enough that it sits on the bookshelves of people like us – those whose day jobs inspire them and who look to their bookshelves for new ideas.

REFERENCES

Ahlgren, E., Azevedo, J. P., Bergmann, J., Brossard, M., Chang, G. C., Chakroun, B., Cloutier, M. H., Mizunoya, S., Reuge, N. & Rogers, H. (2022, 4 January). The global education crisis – even more severe than previously estimated. *World Bank Blogs*. https://blogs.worldbank.org/education/global-education -crisis-even-more-severe-previously-estimated

Aiyar, Y., Davis, V., Govindan, G. & Kapoor, T. (2021). *Rewriting the grammar of the education system: Delhi's education reform (a tale of creative resistance and creative disruption)*. Research on Improving Systems of Education (RISE). https://doi.org/10.35489/BSG-RISE-Misc_2021/01

American University. (2020, 28 July). What is systems thinking in education? Understanding functions and interactions in school systems. https://soeonline .american.edu/blog/systems-thinking-in-education

Beeharry, G. (2021). The pathway to progress on SDG 4 requires the global education architecture to focus on foundational learning and to hold ourselves accountable for achieving it. *International Journal of Educational Development*, *82*, 102375.

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A. & Sandefur, J. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, 1–20.

Borman, G. D., de Boef, W. S., Dirks, F., Gonzalez, Y. S., Subedi, A., Thijssen, M. H. ... & van Berkum, S. (2022). Putting food systems thinking into practice: Integrating agricultural sectors into a multi-level analytical framework. *Global Food Security*, *32*, 100591.

Boulton, J. G., Allen, P. M. & Bowman, C. (2015). *Embracing complexity: Strategic perspectives for an age of turbulence*. Oxford University Press.

Bowman, K., Chettleborough, J., Jeans, H., Rowlands, J. & Whitehead, J. (2015). *Systems thinking. An introduction for Oxfam programme staff*. Oxfam. https://oxfamilibrary.openrepository.com/bitstream/handle/10546/579896/ml -systems-thinking-151020-en.pdf;jsessionid=28872BACEC5FD736EB387C 4A2442F273?sequence=1

Brouwer, I. D., McDermott, J. & Ruben, R. (2020). Food systems everywhere: Improving relevance in practice. *Global Food Security*, *26*, 100398.

Burnett, N. (2019). Invited Essay: It's past time to fix the broken international architecture for education. *International Journal of Educational Development*, *68*, 15–19.

Checkland, P. B. (1981). *Systems thinking, systems practice*. Wiley.

Clinton Foundation. (2019, 19 September). How systems thinking can impact climate change. *Medium*. https://stories.clintonfoundation.org/how-systems -thinking-can-impact-climate-change-14ceaaa7c8ca

CREATE. (n.d.). *Consortium for research on educational access, transitions and equity*. http://www.create-rpc.org

Crouch, L. (2020). *Systems implications for core instructional support lessons from Sobral (Brazil), Puebla (Mexico), and Kenya*. RISE Insight Series. 2020/020. https://riseprogramme.org/sites/default/files/publications/20200708 _RISE%20Insight%202020_20_Crouch_Systems.pdf

de Ree, J., Muralidharan, K., Pradhan, M. & Rogers, H. (2015). *Double for nothing? Experimental evidence on the impact of an unconditional teacher salary increase on student performance in Indonesia* (Working Paper 21806). National Bureau of Economic Research. https://www.nber.org/system/files/ working_papers/w21806/w21806.pdf

de Savigny, D. & Adam, T. (2009). *Systems thinking for health systems strengthening*. WHO.

Dekeyser, K., Rampa, F., D'Alessandro, C. & Molina, P. B. (2020). *The food systems approach in practice: Our guide for sustainable transformation* (Discussion Paper No. 278). ECDPM. https://ecdpm.org/work/the-food -systems-approach-in-practice-our-guide-for-sustainable-transformation

Ericksen, P. J. (2008). Conceptualizing food systems for global environmental change research. *Global Environmental Change*, *18*(1), 234–245.

Evans, D. K. & Mendez Acosta, A. (2021). Education in Africa: What are we learning? *Journal of African Economies*, *30*(1), 13–54. https://doi.org/10 .1093/jae/ejaa009

Faul, M. V. (2014). Future-perfect/present-imperfect: Contemporary global constraints on the implementation of a post-2015 education agenda. *International Journal of Educational Development*, *39*, 12–22.

Faul, M. V. (2016a). Multi-sectoral partnerships and power. Background paper prepared for the *UNRISD Flagship Report: Policy Innovations for Transformative Change*.

Faul, M. V. (2016b). Scoping study: Education systems diagnostics tools. DFID.

Faul, M. V. (2018). Education System Diagnostics: What is an 'education system diagnostic', why might it be useful, and what currently exists? DFID-GPE.

Fenwick, T. (2012). Complexity science and professional learning for collaboration: A critical reconsideration of possibilities and limitations. *Journal of Education and Work*, *25*(1), 141–162.

Flood, R. L. & Jackson, M. C. (1991). *Critical systems thinking*. Wiley.

Fullan, M. (2010). *All systems go: The change imperative for whole system reform*. Corwin Press.

Fullan, M. (2011). Choosing the wrong drivers for whole system reform (No. 24). Centre for Strategic Education Seminar Series. Retrieved on 21 December 2015 from www.cse.edu.au

GEEAP: Global Education Evidence Advisory Panel. (2020). *Cost-effective approaches to improve global learning*. World Bank. https://www.worldbank .org/en/topic/teachingandlearning/publication/cost-effective-approaches-to -improve-global-learning

Gibbs, E., Jones, C., Atkinson, J., Attfield, I., Bronwin, R., Hinton, R., Potter, A. & Savage, L. (2021). Scaling and 'systems thinking' in education: Reflections from UK aid professionals. *Compare: A Journal of Comparative and International Education*, *51*(1), 137–156. https://doi.org/10.1080/03057925 .2020.1784552

Goldstein, J. (1999). Emergence as a construct: History and issues. *Emergence*, *1*(1), 49–72.

GPE: Global Partnership for Education. (2020). *GPE 2025 Strategic Plan*. BOD/2020/11/12 DOC 11. GPE.

Green, D. (2016). *How change happens* (p. 288). Oxford University Press.

Hanson, K. (2015). *What can education systems research learn from health systems research?* (Working Paper 15/003). RISE. https://riseprogramme.org/ sites/default/files/2020-11/RISE_WP-003_Hanson.pdf

Hargreaves, A. P. & Shirley, D. L. (Eds.). (2009). *The fourth way: The inspiring future for educational change.* Corwin Press.

Health Systems Global. (n.d.). *The challenges*. https://healthsystemsglobal.org/ improving-health-systems/the-challenges/

Hickey, S. & Hossain, N. (2019). *Politics of education in developing countries: From schooling to learning*. Oxford University Press.

Honig, D. (2018). *Navigation by judgement. Why and when top down management of foreign aid doesn't work.* Oxford University Press.

Ingram, J. (2011). A food systems approach to researching food security and its interactions with global environmental change. *Food security*, *3*(4), 417–431.

International Food Policy Research Institute. (n.d.). *Food Systems.* https://www .ifpri.org/topic/food-systems

Kaffenberger, M. & Pritchett, L. (2021). A structured model of the dynamics of student learning in developing countries, with applications to policy. *International Journal of Educational Development*, *82*, 102371.

Kaffenberger, M., Savage, L. & Crouch, L. (2021). Schooling without learning: Implications of learning profiles for the global learning crisis. *International Journal of Educational Development*, *86*, SI: Learning profiles.

Laumann, F., von Kügelgen, J., Uehara, T. H. K. & Barahona, M. (2022). Complex interlinkages, key objectives, and nexuses among the Sustainable Development Goals and climate change: A network analysis. *The Lancet Planetary Health*, *6*(5), e422–e430.

Le Nestour, A., Moscoviz, L. & Sandefur, J. (2021). *The long-term decline of school quality in the developing world*. Center for Global Development. https://www.cgdev.org/sites/default/files/consultation-draft-le-nestour-et-al -school-quality-decline.pdf

Lemke, J. L. & Sabelli, N. H. (2008). Complex systems and educational change: Towards a new research agenda. *Educational Philosophy and Theory*, *40*(1), 118–129.

Levy, B., Cameron, R., Hoadley, U. & Naidoo, V. (Eds). (2018). *The Politics and Governance of Basic Education: A Tale of Two South African Provinces*. Oxford University Press.

Lewin, K. M. (2009). Access to education in sub-Saharan Africa: Patterns, problems and possibilities. *Comparative Education*, *45*(2), 151–174.

Little, A. W. & Lewin, K. M. (2011). The policies, politics and progress of access to basic education. *Journal of Education Policy*, *26*(4), 477–482.

Magrath, B., Aslam, M. & Johnson, D. (2019). Raising learning outcomes in the education systems of developing countries: Research designs, methods, and approaches. *Research in Comparative and International Education*, *14*(1), 3–6.

Mason, M. (2008). Complexity theory and the philosophy of education. *Educational Philosophy and Theory*, *40*(1), 4–18.

Mason, M. (2009). Making educational development and change sustainable: Insights from complexity theory. *International Journal of Educational Development*, *29*(2), 117–124.

Meadows, D. (2008). *Thinking in systems*. Chelsea Green Publishing.

Morrison, K. (2008). Educational philosophy and the challenge of complexity theory. *Educational Philosophy and Theory*, *40*(1), 19–34.

Muralidharan, K. & Singh, A. (2020). *Improving public sector management at scale? Experimental evidence on school governance in India* (Working Paper 28129). National Bureau of Economic Research. https://www.nber.org/papers/ w28129

Nature Food. (2020). Systems thinking, systems doing. *Nature Food*, *1*(659).

Ndaruhutse, S., Jones, C. & Riggall, A. (2019). *Why systems thinking is important for the education sector*. Education Development Trust. https:// www.educationdevelopmenttrust.com/EducationDevelopmentTrust/files/17/ 17fec588-e413-461b-a107-78b6569304cc.pdf

Pan, X., Valerdi, R. & Kang, R. (2013). Systems thinking: A comparison between Chinese and Western approaches. *Procedia Computer Science*, *16*, 1027–1035.

Penn GSE. (n.d.). *International Educational Development*. https://www.gse
.upenn.edu/academics/programs/international-educational-development
-masters

Piper, B., Zuilkowski, S. S., Dubeck, M., Jepkemei, E. & King, S. J. (2018).
Identifying the essential ingredients to literacy and numeracy improvement:
Teacher professional development and coaching, student textbooks, and struc-
tured teachers' guides. *World Development*, *106*, 324–336.

Posthumus, H., de Steenhuijsen-Piters, B., Dengerink, J. & Vellema, S. (2018).
Food systems: From concept to practice and vice versa (Wageningen
Economic Research memorandum). Wageningen Economic Research. https://
doi.org/10.18174/464054

Pritchett, L. (2015). *Creating education systems coherent for learning outcomes:
Making the transition from schooling to learning* (Working Paper 15/005).
RISE. https://riseprogramme.org/publications/creating-education-systems
-coherent-learning-outcomes

Ramalingam, B. (2013). *Aid on the edge of chaos: Rethinking international
cooperation in a complex world.* Oxford University Press.

Ramalingam, B., Laric, M. & Primrose, J. (2014). *From best practice to best fit:
Understanding and navigating wicked problems in international development*.
ODI.

Rayner, C. & Bonnici, F. (2021). *The systems work of social change: How
to harness connection, context, and power to cultivate deep and enduring
change*. Oxford University Press.

RISE. (n.d.). *RISE*. https://riseprogramme.org

Rittel, H. W. J. & Webber, M. M. (1973). Dilemmas in a general theory of plan-
ning. *Policy Sciences*, *4*(2), 155–169.

Savage, L. (2013). *Do we tell them what to do or let them choose? Ownership
in the Malawi education sector*. Unpublished PhD thesis, University of
Cambridge.

Steer, L. (2020, 22 October). #SaveOurFuture: Defend Education. *The Education
Commission.* https://educationcommission.org/updates/saveourfuture-defend
-education/?cn-reloaded=1

Steiner-Khamsi, G., Faul, M. V., Hopkins, A. N., Baek, C. & Iwabuchi, K.
(2022). *Strategic review of global and regional evidence and knowledge initi-
atives, networks and platforms in education*. UNESCO.

Stern, J., Piper, B., Jukes, M. & Sitabkhan, Y. (2021). *Learning at scale – A call
for successful large scale numeracy and government led programmes.* Center
for Global Development 2021. https://www.cgdev.org/blog/learning-scale
-call-successful-large-scalenumeracy-and-government-led-programs

Stroh, D. P. (2015). *Systems thinking for social change. A practical guide to
solving complex problems, avoiding unintended consequences, and achieving
lasting results*. Chelsea Green Publishing.

UNESCO. (n.d.). *Futures of Education*. https://en.unesco.org/futuresofeducation/

UNGA: UN General Assembly. (2015). *Transforming our world: The 2030
Agenda for sustainable development. Resolution 70/1*. Adopted by the General
Assembly on 25 September 2015. Doi:10.1007/s13398-014-0173-7.2.

USAID: US Agency for International Development. (2016). *The 5Rs framework in the program cycle.* https://usaidlearninglab.org/sites/default/files/resource/files/5rs_techncial_note_ver_2_1_final.pdf

Wane, W. & Martin, G. H. (2013). *Education and health services in Uganda. Data for results and accountability* (82666). World Bank. https://www.sdindicators.org/sites/sdi/files/countryreports/SDI-Report-Uganda_0.pdf

World Bank. (2018). *Learning to realize education's promise.* https://www.worldbank.org/en/publication/wdr2018#:~:text=at%20the%20center.-,The%202018%20WDR%20explores%20four%20main%20themes%3A%201)%20education's%20promise,make%20systems%20work%20for%20learning

World Bank. (2019). *Ending learning poverty: What will it take?* https://openknowledge.worldbank.org/handle/10986/32553 License: CC BY 3.0 IGO.

World Bank. (2020) *Sustainable Development Goals*. https://datacatalog.worldbank.org/dataset/sustainable-development-goals

Wyse, D. & Torrance, H. (2009). The development and consequences of national curriculum assessment for primary education in England. *Educational Research*, *51*(2), 213–228.

PART I

From practice to systems thinking

2. Reflections on systems practice: implementing teaching at the right level in Zambia

Varja Lipovsek, Laura Poswell, Ashleigh Morrell, Devyani Pershad, Nico Vromant and Abe Grindle

INTRODUCTION

Evidence suggests that if a child does not learn basic literacy and numeracy skills in the early grades of primary school, they get left behind in the higher grades – no matter how many years they spend in school. In middle to upper primary school, most curricula move away from building foundational skills, so children who have not grasped these skills do not have the opportunity to catch up. In many parts of the world, nonproficiency in primary schools is a widespread and systemic problem. Nonproficiency rates are the highest in sub-Saharan Africa, where in 2015, 88% of children (202 million) who are of primary and lower secondary school age were not proficient in reading, and 84% (193 million) were not proficient in mathematics (United Nations, 2019). Central and Southern Asia does not fare much better: 81% of children (241 million) were not proficient in reading, and 76% (228 million) lacked basic mathematical skills (United Nations, 2019).

In the early 2000s, Pratham, one of India's largest education NGOs, developed what has now become known as Teaching at the Right Level (TaRL) to respond to the failure of the education system to ensure that children obtain a solid grasp of basic literacy and numeracy skills before they leave primary education (Pratham, 2022). Rather than attempting to diagnose and fix the multiple failures of the system that have all contributed to this problem (e.g., ambitious curricula, underpaid teachers, lack of supervised on-site training, meagre teaching supplies, etc.), the TaRL

method zeroed in on the classroom practices, with a focus on enabling and motivating teachers to ensure all students attain basic literacy and numeracy skills. At the classroom level, TaRL is a flexible teaching approach that enables teachers to assess children using a simple tool and then group them according to their learning level rather than their age or grade. Each group is taught starting from what the group already knows, using hands-on learner-centred exercises and methodologies. Children move quickly through the groups; for example, children who master letters then move to a group focused on words. Throughout the entire process, teachers assess their pupils' progress through an ongoing, simple measurement of their foundational skill performance rather than relying on end-of-year examinations. The key to the approach is a short feedback loop between hands-on and adaptable teaching methods and rapid tests, which motivate both students and teachers who can see the progress. Several randomised controlled trials (RCTs) conducted by Pratham and the global research centre Abdul Latif Jameel Poverty Action Lab (J-PAL) have demonstrated that when implemented well, the TaRL approach is effective, works at scale and has led to some of the largest effect sizes rigorously measured in the education literature (Poverty Action Lab, 2019).

Given that TaRL can successfully address a problem that is prevalent and persistent in many Global South countries, one way to maximise reach is by embedding TaRL into public education systems, which have both the mandate and scale required for maximum impact. However, decades of development research have demonstrated that powerful forces keep most public systems stable and that even the most effective solutions and the most convincing evidence are not alone sufficient to spur a system to sustainably change how it functions (Andrews, 2013; Teskey, 2019). Translating a policy change into long-term improved systemic practices takes time and depends on a number of different – and often changeable – factors, including formal ones such as regulations and budgets, but also informal relationships, motivations and incentives. Understanding this complex picture and working with a range of actors is the starting point for TaRL Africa, an initiative created by J-PAL and Pratham to improve learning delivered by African education systems. TaRL Africa's approach is to use the robust yet adaptable TaRL methodology as a springboard for dialogue and co-creation with key actors in a given education system, with the goal of not only improving the acquisition of basic skills for millions of children, but also to reorient education systems to a significantly higher order of functioning.

In the current chapter, we examine the approach of TaRL Africa by drawing on systems theory and illustrating this approach through the case study of how TaRL was introduced – and continues to evolve – in the Zambian education system. TaRL Africa does not subscribe to any particular theory and is best described as an effective *systems practice:* guided by a set of principles, adapted to a given context, grounded in the understanding of incentives and power dynamics of the specific education system. In the next section, we outline the core principles of TaRL Africa, linking them to systems change theory. In the third section, we focus on the evolution of the TaRL programme in Zambia. In the final section, we reflect on what we are learning about working with systems in practice.[1]

SYSTEMS CHANGE AND TARL AFRICA

TaRL Africa uses the evidence of a proven and adaptable classroom methodology as an entry point to engage with systems of education in a given context. Collaboration with the key actors both within and outside the government education system is essential throughout the process. TaRL Africa identifies opportunities within the system to focus on children and their learning, works with key actors to co-create a strategy that enables them to diagnose and begin to address that specific problem, and then works with the wider system to evolve the strategy to tackle some of the root causes of the problem over time.

In systems thinking language, the TaRL approach can be said to engage with the major conditions of a system, from the structural (policies, practices, resource flows, etc.), the relational (power dynamics, relationships and connections, etc.), all the way to the transformative (mental models) (Kania, Kramer & Senge, 2018). In other words, TaRL Africa focuses on three high-order systemic levers: information flows, rules of the system, and mental models (Meadows, 1996). However, theory is not at the heart of the approach; there is no blueprint for how these levers are activated (though there are numerous lessons learned), nor is there a pre-set sequence in which the conditions must be addressed. Instead, TaRL Africa is guided by a set of principles focusing on working

---

[1]   The authors of this chapter are involved with TaRL Africa in various ways: from management to implementation and technical support to funding. We acknowledge that the voice of the Government of Zambia is not directly represented in this chapter, and we do not speak on its behalf.

with systems and that orient the evolution of a specific, targeted strategy within each context.

The first orienting principle of the TaRL Africa approach is to *be driven by learning outcomes,* using locally generated data to focus on the agenda. As straightforward as this sounds, in many contexts, the main measures of success in education still revolve around inputs such as enrolment or building classrooms. Although undoubtedly important, TaRL Africa focuses the purpose by keeping the actual learning outcomes of children at the centre. This includes the big picture at the national level, but it is generated from the most granular level: in the classroom, at the level of the individual child. The TaRL approach enables teachers to generate and use data immediately while also equipping them with practical tools and methods to help children catch up in their learning. The data are aggregated upwards through the system to construct the subnational and national pictures of learning.

A second orienting principle of TaRL Africa is to be *pragmatic and deeply grounded in the local context* and realities of systems, always with an eye on what's needed at full scale. This means looking for opportunities to connect with and *build on locally defined priorities*. Although the government remains the backbone of the system, in sub-Saharan African education systems, there are usually a number of influential actors, and working in partnership with them is critical to finding a workable path forward. For example, local civil society actors are often needed to build grassroots support and demand for improved education outcomes; international non-governmental organisations can help the government deliver certain components; and funders are needed for resourcing. What is important here is to centre on those actors rooted in the context who have the mandate to define and address education problems in the long term.

Third, TaRL Africa recognises and addresses both *the technical and human components of the system* from the ground up, starting with the classroom: the children and their teachers. Through TaRL, teachers have simple and effective tools to personally connect with each pupil, assess their learning levels and act on these data in real time by dividing them into learning level groups. They are also equipped with the tools to build the foundational skills of these learners in their respective groups. The teachers move learners to the next group when ready and keep track of progress over time. Teachers master this approach through intensive practice classes, where they strengthen their knowledge and confidence

through learning by doing. They are closely mentored and supported by mentors, who are often government education civil servants.

The data are aggregated upwards, to the school, district, region and national levels, and at each stage this provides a picture of learning achievements and gaps. People at different levels of the system are technically trained to understand these data, but they are also encouraged to interpret them as a real portrait of how the nations' children and teachers are faring. They are encouraged to see their role in identifying gaps to help address them for the benefit of the children, rather than worrying about exposing bad results or not reaching curriculum goals.

Fourth, TaRL Africa works to strengthen the system to the degree that TaRL Africa as an entity is no longer required; in other words, TaRL Africa *plans ahead to its eventual exit.* This means involving governments, other education actors and local civil society from the start, intentionally co-creating throughout the process. The role of TaRL Africa changes while responding to specific systems' needs at different stages in terms of how much the system can take on, evolving support functions and dynamics to respond to the system's evolving needs and ceding control of decisions.

The above core principles of TaRL Africa were developed through the work of practitioners with decades of experience working with both evidence and the systems of education in the Global South. In the following section, we describe the introduction and evolution of TaRL in Zambia's education system. We have chosen Zambia as an example because the work has been ongoing for six years (and still continues), yielding a number of interesting lessons about working with systems. In addition, TaRL Africa supports work in nine additional sub-Saharan African countries,[2] following the same principles but often delivered through a different model and by a different set of actors.

CASE STUDY: TARL IN ZAMBIA

The evolution of TaRL in Zambia can be described in three main phases (Figure 2.1). The first phase focused on co-creating a context-appropriate solution to an urgent need identified by the Ministry of Education (MoE). The second phase revolved around inspiring and enabling the ministry to

---

[2]   Botswana, Côte d'Ivoire, Ghana, Kenya, Madagascar, Mozambique, Niger, Nigeria and Uganda.

expand, from piloting an initiative to integrating it into their functions and aligning other key education actors in support. The third phase, which is still ongoing, focuses on growing the initiative to cover a national scale and ensuring its sustainability through functioning systems and both human and financial resources. The actions, decisions and sequencing within each phase are contextually specific yet align with TaRL Africa's core principles, as described above.
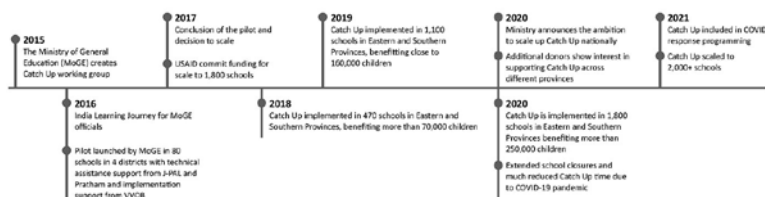
Source: Authors.

Figure 2.1   *Case study: TaRL in Zambia*

Phase One: Co-creating a Context-appropriate Solution to an Urgent Need

In 2015, Zambia was facing a learning crisis. It had been ranked last in measures of literacy and numeracy by the 2011 Southern and East Africa Consortium for Monitoring Education Quality (SACMEQ), and a 2014 national assessment found that 65% of Zambian second-grade learners were unable to read a single word in their local language. The primary school curriculum had recently been revised, and the government and supporting partners were experiencing challenges in rolling it out across the country. There was no shortage of education development aid in Zambia, but the various initiatives were not necessarily aligned and did not adequately address the problem of low basic learning levels. The ministry was alarmed at the low levels of basic literacy and numeracy in the upper primary grades and was looking for ways to address this.

At the same time, J-PAL Africa's policy team was searching for opportunities to work with policymakers to use evidence to address the challenges in their system. The beginnings of this process in Zambia highlight the application of several core principles of *systems practice*. First, there was an organic opportunity where the *system was open to*

change and where TaRL methodology offered a suitable approach. Second, action was driven by the focus on low *learning outcomes* – rather than inputs or processes in the system. Zambia had already implemented several projects that had targeted foundational literacy skills, but most had not shown an improvement in actual learning outcomes. Third, priorities were defined by analysing the *specifics of the problem as manifested in the Zambian context*; it emerged that there was no basic skills safety net for children in grades 3–5, so there was a clear gap where this massive problem existed.

Following consultations with J-PAL Africa on the evidence related to improving foundational skills, the MoE's Office of Standards and Curriculum was keen to adopt the TaRL approach, providing *internal leadership and championing* the idea within the ministry and also with a number of development agencies. J-PAL Africa further supported this process by engaging with key development actors, sharing evidence from other contexts and discussing how it could be helpfully applied in the Zambian context. Critically, the J-PAL Africa team believed that the entire process needed to be owned and managed by the key system actor, in this case, the MoE, with J-PAL and Pratham providing insights and technical assistance, not driving the decision making. With the MoE firmly in charge, an important first move was for the ministry to organise a working group of main education partners, including UNICEF, the British Department for International Development (DfID; now FDCO), the British Council and Innovations for Poverty Action Zambia. This group, under MoE leadership, hatched the plans for a pilot of what the ministry branded the Catch Up programme. The funds for the 80-school pilot were also sourced by the ministry from the Global Partnership for Education, with additional support from UNICEF and J-PAL's Government Partnership Initiative (GPI). The ministry also partnered with VVOB–Education for Development for additional mentoring and monitoring of implementation support during the pilot.

The primary goal of J-PAL Africa was to equip the MoE with the evidence they needed to make informed decisions, building on the policy lessons from the randomised evaluations. Although the TaRL intervention has been robustly proven to work across contexts, not all implementation models lead to equal improvements in learning outcomes. For example, volunteer-led models have been shown to lead to the quickest learning gains but are hard to scale and sustain. Teacher-led models show a smaller impact per child on learning outcomes, but they are more scalable and sustainable. The MoE chose to test two different teacher-led

delivery models during the pilot: a TaRL holiday camp and an hour a day of additional TaRL instruction during term time. In both models, the MoE built on the evidence generated through research showing that for a teacher-delivered TaRL model to be effective, it requires frequent visits from mentors and dedicated time for TaRL classes (Banerjee et al., 2016).

Another key principle of *systems practice* is working with the *people within the system*. This means ensuring that the actors at all levels of government connect with the problem but also with the idea that progress is possible within the existing parameters. Critical to sparking enthusiasm and commitment – but most importantly hope – was a learning journey by two key ministry officials to see TaRL in action in India and introduce them to Pratham. As a result, Pratham created a team of four staff members to support Zambia. This team spent much time in Zambia working with ministry officials to codevelop a TaRL approach that makes sense for the country, with materials in relevant local languages and contextualised teaching activities delivered within the Zambian education system. They also trained ministry officials and provided remote support. VVOB was MoE's in-country partner of choice to complement these efforts by providing additional mentoring-monitoring boots on the ground. Their mission alignment to improve quality education through building government capacity and their existing relationships and location working from within the MoE meant they could enhance programme delivery and help build quality checks into the system.

The pilot was implemented in 80 schools in 2016 and proved significant in a number of ways. First, it confirmed that most third–fifth-grade learners in Zambia lacked basic reading and mathematics skills, with more than half of children in grades 3 to 5 in the Catch Up pilot schools at baseline unable to read words. Second, the process evaluation results showed that the programme was well implemented, which is essential in establishing its feasibility. Internal monitoring largely occurred as planned; teachers stuck to the key principles of TaRL, and they continued to implement the programme over time. Third and most important, learning outcomes improved markedly during the one-year pilot period. According to government data, the share of children who could not even read a letter fell by 26 percentage points, from 33% to 8% during the pilot period, and the share of children reading with basic proficiency (a simple paragraph or a story) grew by 18 percentage points, from 34% to 52%. In arithmetic, the share of students in the beginner group (who could not even complete two-digit addition sums) fell by 16 percentage points,

from 44% to 28%, and the share of students with basic proficiency (able to complete two-digit subtraction) rose by 18 percentage points, from 32% to 50%.

The successful pilot proved to be a *tipping point* within the MoE: the ministry decided to expand the programme based not only on positive results, but also on the feasibility of implementation because the pilot had been executed with and through the government's own systems. The MoE had chosen to pilot two different TaRL models, ultimately selecting the one they believed would be the most sustainable in their system. With this, the MoE generated enthusiasm among other development partners, and the Catch Up programme was awarded a grant by USAID Zambia (through the USAID Development Innovation Ventures structure) to expand the programme to reach 1,800 schools over three years. J-PAL Africa hired a small in-country team to support programme development and management; together with the MoE, J-PAL selected VVOB to provide on-the-ground mentoring and monitoring resources. The Pratham team visited frequently and provided remote support. J-PAL, Pratham and VVOB have continued to work in close partnership with the MoE and other education stakeholders in Zambia as Catch Up has developed and grown.

Phase Two: Enabling the Ministry to Expand from Piloting to Integration

The way TaRL began in Zambia laid the foundation for a government-owned programme. However, while reaching the 'tipping point' of change is necessary as a starting point of a deeper systems-level reform, the trajectory of that change depends on the continuous collaborative pursuit of a scalable and effective approach requiring ongoing reflection and adaptation. Following another core principle of working with systems, J-PAL and Pratham *adapted the support provided* to the MoE to keep up with the MoE's own evolving capacities and needs. The support model transformed from crafting and improving the localised design of Catch Up and the direct mentoring of teachers at schools to working with zone, district and provincial leaders of the programme in delivering the programme's core functions. The real 'magic' of *systems practice* happened in this transformation from a pilot into institutionalised changes in key functions of core systemic actors. The success can be seen in the government data showing year-on-year improvements in learning outcome progress since and including the 2016 pilot.

In 2019, the MoE's roll-out of the Catch Up programme reached approximately 1,100 schools, and in 2020 it was implemented in 1,800 of the approximately 10,000 government primary schools (Figure 2.2). Scaling did not mean simply replicating the pilot in more schools. Instead, the focus was on iterating the model to build on the operational lessons from the pilot, here with the goal of continually improving the core approach to meet the context and further ingraining the approach into the various levels of the education system, as well as establishing ownership and links across the levels.
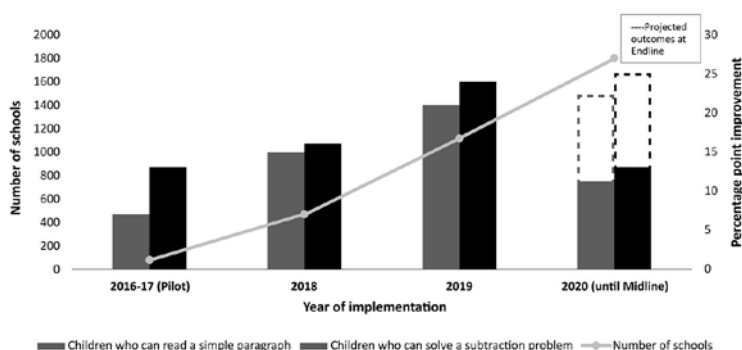
Source: Authors.

Figure 2.2   Catch Up Zambia: scale vs. outcomes

J-PAL and Pratham formalised their partnership in Africa as 'TaRL Africa' and, together with the MoE and VVOB, focused on *the incentives and motivations of key human resources* – the teachers. Critically, the Catch Up programme does not treat teachers as the 'end of the line' of the long human resources chain. Rather, teachers are seen as the starting point: if the TaRL approach does not work for teachers, it will not be successful. An essential part of the Catch Up approach is participatory methods meant to create an atmosphere of relaxed and fun learning, including singing and games. In the initial stages, not all the teachers even within the same school reacted in the same way. Some were very enthusiastic and started implementing the methodology. Some were not at all interested. Most took a neutral stance because they had seen multiple projects come and go and did not want to invest much time and energy. However a critical mass of teachers in the selected districts

did implement it, and because Catch Up generates outcome data very quickly, these teachers were energised by the clear improvements in learning among their pupils. Through the information flow architecture set up for TaRL, the data reached district education officials quickly as well. Recognising the promise of the approach, district officials sought to motivate other schools and teachers to use Catch Up, for instance by sending out letters signalling that this was a MoE project supported by MoE partners. Partly as a result of this external motivation by district officials and partly as demonstrated by the enthusiasm of fellow teachers, the large group of 'neutral' teachers in the selected districts started implementing Catch Up. TaRL Africa recently completed a systematic qualitative study aiming to understand what drives Zambian teachers to change their mode of instruction; this work shows a clear picture of reported behaviour change, with the change being attributed to the Catch Up programme strictly dominating as relative to other programmes (de Barros, Henry & Mathenge, 2021). The change related to Catch Up is most commonly noted as an increased understanding of learners' needs, adoption of the Catch Up methodology, increased interaction with learners and increased use of teaching materials. Teachers also frequently mentioned that they applied the Catch Up methodology in other classes (which are conducted during regular school hours and not targeted by the Catch Up programme).

If one thinks of teachers and schools as being at the centre of the public education system in Zambia, then the district- and province-level officials are the next layer. Understanding and addressing their needs, incentives and motivations is another key part of Catch Up. For example, the education civil servants from provinces and districts were selected as master trainers from the outset so that the system would be trained by its own officials. How important district officials were to the implementation of Catch Up became clear as short-term measures of learning outcomes fluctuated with the natural turnover and shifting of relevant district staff. When a district official from a well-performing district moved to a poorly performing district, that district started improving; the opposite was true as well. Recognising this, Catch Up officials, supported by TaRL Africa and VVOB, worked with provincial education officers and developed a plan for how to engage district officials. For instance, continuing the South-to-South collaboration, Pratham staff came to Zambia to support the ten-day training of these government officials and then returned to help train a larger contingent to lead an expanded scale-up. In addition, in each year of the roll-out, education staff from the provincial, district

and school levels were engaged in regular working groups that met three times a year to improve different aspects of the programme – including the teacher guides, lesson procedures, classroom activities and measurement processes – but also very contextual and practical problems that could derail the effort if not addressed. For instance, province and district officers originally had a heavy load of school visits to deliver monitoring, mentoring and coaching. However, because of logistical and transport issues, they were not able to fully carry out these functions. The Catch Up team shifted the model: the functions of mentoring and coaching were allocated to senior teachers, who were also trained and supported to deliver these functions. The role of province and district officers shifted more towards accountability: sending instructions to schools, supporting schools as necessary, organising reflection meetings, monitoring the set-up of the assessments and acting upon assessment results.

Another important part of the human–technical link was the further strengthening of the monitoring and feedback system, with more data collection and analysis conducted at the school level by senior teachers. Data collection and aggregation was not easy for many of the schools and officials. The Catch Up team had originally placed this responsibility with the district and province officers of the MoE's Directorate of Teacher Education and Specialised Services, but during implementation it became clear that it was in fact the MoE's Directorate of Planning and Information that had the mandate and capacity to collect education data. Subsequently, this directorate was trained in Catch Up and began assisting schools and districts. This encouraged data-based decision making to take place at school and made aggregation and analysis easier at the zone and district levels. In fact, the classroom-based assessments and use of data at all levels – from classroom, to school, to district, to province, to national – played a critical role throughout this phase. For the first time, everyone within the education system had easy-to-understand and reliable data generated through their own systems. This created not only visibility and insights into performance at every level, but also ownership. In this way, the data could give stakeholders a trusted and tangible resource: the data could help identify problems and track progress, as well as show and celebrate the results.

Phase Three: Grow, Adapt and Sustain

From the start, TaRL Africa's vision was to enable ministry leadership to take charge of improving basic education outcomes supported by local

actors and make the best use of the development aid and international actors present in Zambia. Every component of the programme was designed together with the MoE – the delivery model, the materials and activities, and the monitoring and measurement system. This true collaboration enables the models to be owned by the government, ingrained in practice and, therefore, hopefully sustained over a long time. A signal of the usefulness of TaRL to the overall system became clear through the COVID-19 pandemic. COVID-19 had a significant negative effect on primary education in sub-Saharan Africa (Dang et al., 2021). In Zambia, schools were closed for six months in 2020 and for two months in 2021. In response, the MoE made Catch Up part of its 'Education Contingency Plan for Novel Coronavirus (COVID-19)', and Zambia's eastern province decided that Catch Up would become the main method to work on learning loss and foundational skills. Catch Up was also taken up in the COVID-19 Global Partnership for Education emergency project, which covers 20 districts, while the LEGO Foundation provided COVID-19 grant support to extend Catch Up coverage to the Lusaka Province.

As the system has grown stronger, the support provided by TaRL Africa and VVOB has further evolved to address long-term planning and management needs within the system. Specifically relevant to TaRL, the ministry is planning to integrate selected activities that comprise management of the Catch Up programme into job descriptions of key ministry actors at the national, regional and district levels and to incorporate the TaRL methodology into pre-service and post-service teacher training. The focus of TaRL Africa and VVOB is shifting away from the Catch Up initiative itself, to more generally supporting the MoE to set its own vision and strategy and manage its execution across a range of programmes and actors.

As the Catch Up programme delivered impressive results and generated visible enthusiasm from the education sector, many important education donors in Zambia became increasingly keen to fund it. This support is essential because the Zambian government has limited resources and needs the collaboration of international actors. At the same time, it means there is the risk of fragmenting the focus by well-meaning but not necessarily aligned actors, unless the MoE can hold on to key oversight and management functions, as well as setting the overall strategic direction. For example, a year after the introduction of Catch Up, USAID began the Let's Read literacy project, which targets learners in Early Childhood Education (ECE) and grades 1, 2 and 3 in five Zambian provinces, of which two were also the first Catch Up provinces. Not

surprisingly, the parallel programmes competed for scarce government resources, particularly time of MoE staff and time among education civil servants to help organise and participate in training. The MoE is managing the two programmes by aligning their complementarities, from the classroom level all the way to national level. The USAID's Let's Read provides teachers and learners with a multitude of materials that are helpful in delivering the primary literacy programme in ECE and grades 1 to 3. Catch Up is primarily remedial in nature; it focuses on improving foundational literacy and numeracy skills of those learners in grades 3 to 5 who had fallen behind the regular curriculum. From the MoE's perspective, it made sense to welcome both initiatives, and in this case the complementarities are clear. However, the universe of influential actors in the Zambian education system is large and maintaining a unifying vision and strategy while balancing the sometimes competing priorities of various actors can be a challenging task. TaRL Africa and VVOB's role is now further evolving to work with the MoE on developing their own vision and strategy, and on strengthening central coordination.

TaRL Africa's foundational work in Zambia and the lessons learned on this journey have sparked interest in the TaRL approach across the continent, giving rise to many TaRL-inspired programmes and leading to the formation of the TaRL Africa organisation to provide this type of support across the region. At the same time, TaRL Africa, Pratham and J-PAL do not believe that the TaRL approach is the 'silver bullet' that can fix foundational learning problems in all contexts. Although TaRL is indeed a specific methodology, the TaRL Africa approach is much more about understanding the core of the learning problem in a specific context and figuring out whether the TaRL methodology might be part of the pathway of improving the functioning of the education system. In Zambia itself, TaRL Africa continues to learn and adapt, acknowledging that there is still some way to go to achieve the country's learning outcomes goals and those of national scale and full system ownership. The work is ongoing, and so are the lessons being learned.

REFLECTIONS ON WORKING WITH SYSTEMS OF EDUCATION

Perhaps the best way to describe TaRL Africa is to say that it dances with the systems it works with; guided by principles, it adapts and adjusts according to the context, actors, boundaries and incentives (Meadows, 2001). The overall goal to improve the delivery and outcomes of educa-

tion remains the same, but the methods and sequence can – and do – vary. The focus is on government education institutions because they have the mandate to reach primary school students at scale, but the process usually involves collaboration with a range of different educational actors, including civil society and international donors. There is no pre-set menu of who must be involved; the constellation depends on the combination of interest, opportunity and capacity, and it can change over time. In Zambia, the opportunity arose within the Ministry of Education, in large part because the ministry itself had already identified remedial teaching as a priority to shore up basic learning levels in primary school. There was genuine organic interest from a system actor in an influential enough position to make TaRL a national project, drawing in the support of a number of important non-government actors, such as UNICEF, GPE and others.

To work this way, TaRL Africa must be comfortable with uncertainty, and it must be able to change course, as needed. Rigid recipes or must-follow maps would be useless. Instead, TaRL Africa relies on its core principles to help it navigate each particular context. In turn, these principles help unlock some of the key levers in education systems.

The focus on data and teachers is an ingenious backbone of TaRL, through which it works on one of the highest levers in a system – the mental models. Teachers are a critical part of the civil service and account for a significant proportion of national spending in sub-Saharan African countries (Crawfurd, 2020). At the same time, research has consistently revealed serious challenges in the profession and how it delivers education, frequently marked by low motivation and high absenteeism among teachers (Bold et al., 2017). However, the picture is complicated. Teacher advocates often point to systemic malfunctions, such as inadequate training, low or inconsistent pay and overambitious curricula, as the reason for the failure of teachers to show up and deliver quality teaching. The solutions are not straightforward, and there are many examples of high-level reforms (e.g., the revision of curricula or increases in teacher salary) that have yielded no impact on learning (de Ree et al., 2017). It is quite common for a ministry of education and associations of teachers to be stuck at an impasse, blaming one another for poor performance. The TaRL approach can cut through this by challenging how teachers see themselves and how the system perceives them. It taps into the human motivation and potential of teachers by giving them autonomy to focus on the children in their classroom: test their learning levels, group them accordingly, teach through interactive, simple but effective means, re-test

to observe the results and adapt the methods accordingly. The teachers are no longer the last cogs in the machine passively delivering content; they become central figures with the most visibility into and capacity to respond to children's learning needs.

Through this focus on granular, usable data that can be aggregated up to regional and national results, TaRL influences another powerful system lever: the generation, flow and access of critical information. Along the way, some rules of the system may become redefined as well. At each important node (classroom, school, district, province, etc.), the TaRL Africa approach enables the actors to see and feel the data, to have a sense that it is grounded in their schools and their children and to take action at their level, which will further support learning. Feedback loops are built along the trajectory because each level communicates with other levels. This is in stark contrast to periodic high-stakes examinations that are set according to ambitious curricula (and, therefore, usually out of tune with much of the student body), often designed by experts with little or no input from those delivering education, administered through top-down protocols, and analysed centrally. TaRL does not replace examinations. It has a different purpose: to generate trusted data on learning that are useful immediately and directly to the teachers administering the simple tests while strengthening the monitoring mechanisms to allow these data to be shared and used up and down the system. In other words, implementing TaRL means wrestling some of the power from the upper echelons of the system (where the curricula are set and examinations are designed and administered) and distributing it across the system.

Likely one of the least surprising lessons is that this process – particularly when the redistribution of power is in question – is often contested, does not follow a straight 'progress' line and takes longer than originally planned. In fact, although planning is essential, commitment to co-creation and local ownership means responding to needs as they appear and working at a pace that builds strength in different areas and actors of the system over time. In turn, the system's capacity to absorb change and deliver new components also adjusts, further redefining the needs and pace.

In this chapter, we have attempted to describe the practice of working with systems of education through a case study of the evolution of the TaRL initiative in Zambia. Although it is very tempting to package this into a series of 'best practices' that could be replicated elsewhere, we believe that the primary lesson our case study illustrates is that successful work with systems is guided by a series of principles and a clear purpose,

while the choice of actions and processes is thoroughly adaptive over time. We – the authors and organisations we represent – have also learned going through the process and are influenced by the progression of the Catch Up programme and the ongoing transformation of the Zambian education system. We offer some of our reflections below.

At Pratham, working with governments and communities is at the core of our approach to partnership. We bring know-how and expertise, but each journey is one of discovery, collaboration and co-creation. The Zambia experience taught us how to do this in a different context and culture. When starting the collaboration with the Zambian MoE, we were not aware of the availability of resources in the government system to design and execute an innovation like TaRL. However, the involvement and motivation of government officials from the very start has allowed us to build a programme well-suited to the realities of the Zambian system. We have honed our approach to partnerships, focusing on the flexibility and openness to adapt and adjust based on the experiences and expertise of our partners. We look forward to continuing this journey through our TaRL Africa and global partners and making improved learning a possibility for children across the world.

At VVOB–*Education for Development*, we partner with ministries of education to provide innovative solutions that ensure education systems are equitable and inclusive, and provide quality education. The Catch Up experience has encouraged us to reflect more about the conditions, assumptions and processes that enable an education system to bring the lesson from a small pilot into core functions of the system to reach a national scale. The added focus on quick data collection has helped us further fine-tune our capacity development strategy towards the MoE, its staff, headteachers and teachers. The trajectory of the initiative overall has underscored the power of networking with other organisations. VVOB is eager to further assist the MoE in its endeavour to bring Catch Up to all schools in Zambia. Building on the experiences in Zambia, VVOB has started piloting TaRL in Uganda and is eager to learn more about which model works best in which context and to contribute to a continental learning and research agenda on TaRL.

At J-PAL Africa, we focus on supporting policymakers and practitioners in using evidence to improve lives. The Catch Up journey has stretched our understanding of what bridging the gap between evidence and action means. We have experienced the value and importance of deeply engaging in the implementation details of a programme to ensure that the lessons from the evidence are carefully operationalised. We have

actively incorporated complementary research methods to ensure the boundaries of what specific evaluations measure are expanded to include critical global lessons, theory, on-the-ground experience and local contextual considerations. Our work within Catch Up has been pivotal in our development of more general frameworks to inform our broader efforts to move from research to policy to practice. Most of all, the Catch Up and TaRL Africa journey more generally have affirmed our belief in the power of collaboration. Catch Up's beginning and growth have required the contributions of many actors with diverse skill sets. These partnerships are not always simple to navigate, but they have been grounded in the common core value of improving learning outcomes for children in Zambia and doing so sustainably into the future. It has been encouraging to see what we can achieve together.

At Co-Impact, we believe in supporting organisations like TaRL Africa, which are squarely focused on working with systems to foster enduring improvements in how they function, and we try to structure our support in ways that best enable such work. A key aspect to TaRL Africa's success is the team's ability to maintain a laser focus on the long-term goal while having the freedom to manage its own choices regarding how it engages with and supports the system in any given context. In fact, across our grants we see that organisations are successful system actors precisely because they can adapt and make choices, not despite this. What is critical is alignment of purpose, which we state as improved development outcomes for people *through* systems change. As funders, this means that we need to provide TaRL Africa and others with long-term funding that meaningfully contributes to the full cost of the effort (including the critical costs needed to lead, manage and further strengthen a sophisticated organisation), that is fixed on a core set of long-term results (such as enduring improvements in the system that leads to better results for millions of students) but that is also flexible and adaptive, giving TaRL the freedom to truly 'dance' with the system and respond well to its dynamics and changes in the context.

TaRL Africa envisions a future in which all children who are in school learn; it believes that the optimal way to arrive in that future is to enable systems with the mandate and the reach to perform better.

We close with a quote from Donella Meadows (2001):

> The future can't be predicted, but it can be envisioned and brought lovingly into being. Systems can't be controlled, but they can be designed and redesigned. We can't surge forward with certainty into a world of no surprises, but

we can expect surprises and learn from them and even profit from them. We can't impose our will upon a system. We can listen to what the system tells us, and discover how its properties and our values can work together to bring forth something much better than could ever be produced by our will alone. We can't control systems or figure them out. But we can dance with them!

REFERENCES

Andrews, M. (2013). *The limits of institutional reform in development.* Cambridge University Press.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji. S., Shotland, M. & Walton, M. (2016). *Mainstreaming an effective intervention: Evidence from randomized evaluations of 'Teaching at the Right Level' in India.* No. w22746. National Bureau of Economic Research.

de Barros, A., Henry, J. & Mathenge, J. (2021). What drives teachers to change their instruction? A mixed-methods study from Zambia. Unpublished manuscript. Cambridge, MA.

Bold, T., Filmer, D., Martin, G., Molina, E., Stacy, B., Rockmore, C., Svensson, J. & Wane, W. (2017). Enrollment without learning: Teacher effort, knowledge, and skill in primary schools in Africa. *Journal of Economic Perspectives*, 31, 185–204. 10.1257/jep.31.4.185.

Crawfurd, L. (2020). *Everything you ever wanted to know about teacher labor markets*. Center for Global Development. https://www.cgdev.org/blog/everything-you-ever-wanted-know-about-teacher-labor-markets

Dang, H. A., Oseni, G., Zezza, A. & Abanokova, K. (2021). *Impact of COVID-19 on learning: Evidence from six sub-Saharan African countries.* World Bank. https://openknowledge.worldbank.org/handle/10986/35676

Kania, J., Kramer, M. & Senge, P. (2018). *The water of systems change*. FSG. https://www.fsg.org/resource/water_of_systems_change/

Meadows, D. H. (1996). Leverage points: Places to intervene in a system. https://donellameadows.org/archives/leverage-points-places-to-intervene-in-a-system/

Meadows, D. (2001). Dancing with systems. *Whole Earth*, *106*, 58–63. https://donellameadows.org/archives/dancing-with-systems/

Poverty Action Lab. (2019). *Teaching at the right level to improve learning*. Retrieved from https://www.povertyactionlab.org/case-study/teaching-right-level-improve-learning

Pratham. (2022). *Teaching at the right level*. Retrieved from https://www.pratham.org/

de Ree, J., Muralidharan, K., Pradhan, M. & Rogers, H. (2017). Double for nothing?: Experimental evidence on an unconditional teacher salary increase in Indonesia. Policy Research Working Paper 8264. World Bank. https://openknowledge.worldbank.org/handle/10986/28987

Teskey, G. (2019). Afterword. In J. T. Seidel & J. Faustino (Eds), *Thinking and working politically in development: Coalitions for change in the Philippines*. The Asia Foundation.

UN. (2019). *United Nations Sustainable Development Goals Overview*. https://unstats.un.org/sdgs/report/2019/goal-04/

3. Collaborative professionalism and education system change: new evidence from Kenya, India and Rwanda

Charlotte Jones, John Rutayisire, Donvan Amenya, Jean-Pierre Mugiraneza and Katie Godwin

> We sit together, and share what we can't do alone.
> Leader of Learning, Rwanda

OVERVIEW

This chapter offers new empirical evidence on the characteristics of collaborative professional relationships, and how these relationships can function to bring about positive education system change. It argues that collaborative professionalism is one example of systems thinking that can be applied to the field of education development, and to large-scale change management in the Global South. By collaborative professionalism, we refer to peer professionals – in this case teachers or headteachers – working together in network structures, such as communities of practice (CoPs), to share learnings and improve their practice. By education system, we mean the people, processes, relationships, resources and institutions that interact to deliver education outcomes (DFID, 2018). The chapter provides a synthesis of emerging empirical evidence from three settings – in Kenya, India and Rwanda – to illustrate the characteristics of collaborative professionalism in contexts where it has been applied at scale to bring about shifts in learning outcomes.

Leading academics recognise that public sector reforms often fail to translate policy into practice, and they call for a better understanding of 'how change happens'. In other words, there has been a push for

a better understanding of large-scale delivery, people and institutional change. For example, Banerjee et al. (2017) talk about 'getting inside the machine' of government to successfully deliver interventions at scale. Pritchett et al. (2010) call for a focus on understanding the black box of delivery and the functioning of public systems and institutions. In education this field, which we refer to as change management, is small and much of the emerging literature focuses on accountability relationships or incentives for change.

We argue that the principles of collaborative professionalism have an essential role to play in this nascent discipline of change management in education development and are pertinent for addressing those complex challenges that require a more systemic solution, such as quality improvement at scale. Although there is established literature in high-income contexts (Hargreaves & Fullan, 2012; Hattie, 2016), there is less evidence from lower-income settings. Through the case studies, we illustrate the shifting relationships between education system actors as change happens. We argue that a positive dynamic for change is not only based on accountability relationships. When relationships are instead based on trust, dialogue, inquiry, peer accountability and a shared purpose, positive change can emerge. A collective understanding of goals can be built, which drives improved problem-solving, more effective practices and better alignment of efforts. We also offer evidence of the type of leadership needed by actors in an education bureaucracy for this collaborative professionalism to flourish, arguing for a strong 'middle' layer of leadership, including change agents and distributed leadership structures, that catalyses a collective spirit.

Together with insights from the literature, this evidence is used to offer a conceptual framework for high-functioning collaborative networks. The chapter draws on EDT research from collaborative professionalism that has been conducted over four years, in partnership with IIEP-UNESCO, the Education Commission, the Foreign and Commonwealth Development Office (FCDO), World Innovation Summit for Education (WISE) and STiR Education.

INTERNATIONAL EVIDENCE

Systems thinking has a long history that spans diverse disciplines and has influenced reform approaches in the public sector. Systems approaches – which aim to address complex problems where the relationships between cause and effect are nonlinear and often unpredictable – are highly

relevant to education, which is characterised by contextual variance and many unique transactions and objectives (school leader to teacher, teacher to student) that can result in compounding and contradictory goals throughout the system (OECD, 2017).

The application of systems thinking to education has only recently emerged, and there are limited theoretical and empirical frameworks guiding research, especially in low- and middle-income contexts. One framework that has been applied in developing countries is the Research on Improving Systems of Education Programme's 5x5 framework, which defines coherence in terms of the alignment of principal–agent flows of accountability across the relationships between actors in a system and their design elements (Pritchett, 2015).

Wider systems thinking looks beyond accountability and identifies the need to address the 'soft issues' of systems relationships – such as trust and commitment – that influence decision making. It recognises the need to bring coherence to the divergent stakeholder assumptions and objectives and to the distributed decision making which is inherent in complex systems. Some research focuses on achieving such coherence by building a shared understanding of the problems and objectives in a system through collective interrogation, debate, dialogue and feedback loops (Senge, 1990; Checkland & Poulter, 2006; Fullan & Quinn, 2016).

A growing body of evidence suggests that professional collaboration succeeds in improving coherence and responding to complexity, by fostering and leveraging collective understanding and action. Professional collaboration draws on soft systems thinking which assumes that actors' conversations and sensemaking processes are vital to constructing and managing systems. Given this assumption, building a collective understanding of problems and objectives can help solve issues not addressed by accountability flows. Public sector theorists have long proposed that thinking about the collective and interrelated nature of public service delivery has particular relevance for addressing complex problems and longer-term goals (OECD, 2017) such as education quality improvement as opposed to command-and-control systems, which are more applicable to predictable or logistical challenges.

In education, at the state level, several studies cite professional collaboration as a characteristic associated with strong student performance (Jensen et al., 2016, McAleavy et al., 2018). At the practitioner level, evidence on the impact of the collective – over individual actors – on outcomes has been well established since Bandura's work in the 1990s. Hattie (2016) ranks teacher collective efficacy – the shared belief in

a group's ability to influence outcomes – as the first of 250 factors influencing student achievement.

A common form of collaboration involves professional learning communities (PLCs), which have been linked to school improvement (Darling-Hammond et al., 2017). In PLCs, teachers enquire together into student learning and how to improve their practice and then implement what they have learned. A related concept is the CoP, which is a form of network that leverages a group's agency to deepen its knowledge around a shared practice (Ianquinto et al., 2011).

However, setting up collaborative structures is not enough; Hargreaves and Fullan (2012) warn against 'administrative contrivances' where the right structures are in place but dynamics of the group fail to support professional improvement. These ideas about effective professional collaboration and group dynamics are based on particular ontological understandings of how actors interrelate, how information flows and how decisions are made. They place an emphasis on the lateral exchange of information and tacit knowledge, shared practice development, peer accountability and fostering intrinsic motivation and agency (Darling-Hammond et al., 2017; Eells, 2011; Hildreth & Kimble, 2004; Priestley et al., 2015).

A body of work develops these ideas further, recognising that improvement is intimately tied to the dynamics of professionals as a group, not simply the sum of individual contributions (Bandura, 1997). Much of this work focuses on building high-functioning collaborative systems. Stoll and Temperley (2009) suggest that group dynamics can be cultivated through a strong focus on the adaptation of professional practices to better serve student outcomes, with all players learning, unlearning and relearning without fear of failure.

Despite overlap and multiple meanings within this family of concepts, a growing body of work broadly labels these collaborative bottom-up approaches to change as 'Model Y', which emphasises the distributed nature of knowledge and collective agency for improvement, contrasting with 'Model X' theories that draw on the tradition of new public management (Table 3.1).

How can these bottom-up professional networks be harnessed to deliver system-wide outcomes? It is increasingly recognised that these Model Y-type approaches alone are not adequate for large-scale change. Challenges have been identified, including uncertainty about the ownership of responsibility, and some roles lacking legitimacy or the capacity

Table 3.1   *Approaches to change management across public sector systems*

Model X	Model Y
Top down, centrally implemented	Locally owned
Delivery chains and structures	Networks, partnerships, ecosystems
Vertical reporting	Lateral connections
Command and control	Empower and enable
Driving performance	Solving problems, cycles of continuous improvement
Principal–agent; beneficiaries	Change agents, co-constructors
Leadership	Trust; relationships; stewardship, facilitation
Goal alignment through incentives and sanctions	Goal alignment through shared values, norms and intrinsic motivation

Source: Authors, building on Honig (2020); Williams et al. (2020); Silvia (2011).

to operate under new forms of autonomy (Wohlstetter et al., 2014; Bodily et al., 2004).

Neither Model X nor Model Y alone offers a clear framework for at-scale change. In recent years, educational thought leaders have offered a strong alternative vision. Hargreaves and Fullan (2012) discuss harnessing collaborative action and shared knowledge for impact at scale, or what they call social capital. They propose 'collaborative professionalism' as a way to move the development of capital beyond groups of teachers to the transformation of the system. Hargreaves and O'Connor (2018) position collaborative professionalism at the intersection of Model X and Model Y. Similarly, Bryk et al. (2015) suggest that many complex problems embedded within education systems require coordinated, collective action beyond practitioners, which involves diverse sources of expertise that can be harnessed through network improvement communities.

Emerging evidence suggests several critical factors for realising these visions of collaborative professionalism. Effective collaboration is founded on relational trust across all stakeholders (Barletta et al., 2017), with group coherence on expectations and obligations (Bryk & Schneider, 2002). Without this, collaboration can be forced and unproductive, what Hargreaves and Fullan (2012) call 'contrived collegiality'.

Effective professional collaboration is intentional and aligned around a shared vision. Hargreaves and Fullan (2012) argue that 'deliberate

design' is needed to support both informal and formal collaboration, and other scholars find that a shared vision or goal is a key factor in positive network outcomes (Barletta et al., 2017; Riley & Stoll, 2004).

Collaborative enquiry can provide a structure for building relational trust and a shared vision, and models of collaborative teacher enquiry have been shown to enhance collective efficacy (Donahoo et al., 2018). This happens because teachers' beliefs change when they reconcile the discrepancies between their initial thinking and assumptions with new ideas that are created through collective enquiry and reflection (Donahoo & Velasco, 2016). Collective reflection can increase shared accountability as individuals allow their beliefs and practices to be influenced by other actors and choose to value collective results over individual ones (Elmore, 2008).

Positive collaborative networks also require access to expertise, with one study (Proger et al., 2017) finding that distinct types of expertise are required to ensure accuracy and build legitimacy for shared work. Similarly, Elmore (2005) argues that distributed expertise within an organisation and the effective use of external expertise are essential elements of improvement. No one person or role can take on all the complex skills or knowledge required to provide education, and Elmore sees improvement as a continuous process that requires different combinations of knowledge and skills from different actors and leaders.

Finally, a culture of trust aligned with a shared vision requires strong distributed leadership. Evidence shows that leadership from a broad range of roles and that emerge at different system levels can support deep collaborative learning (Dumont et al., 2010) and drive shared decision making (Katz & Earl, 2010). Harris (2012) emphasises that distributed leadership is not an alternative to 'top-down' Model X-style management, but it encompasses both formal and informal leaders and acknowledges the interdependencies that shape leadership practice (Spillane, 2006; Cullen-Lester & Yammarino, 2016). Research suggests that collaborative leadership capacity across networks is key to making a sustained impact on school improvement practices and outcomes at scale. Drawing on a range of evidence, Silvia (2011) argues that certain leadership behaviours enable members to harness the knowledge and skills across the group, such as empowering members, developing a common language, fostering trust, openness and inclusiveness and bringing new perspectives.

ABOUT THE INTERVENTIONS: COLLABORATIVE PROFESSIONALISM IN THREE CONTEXTS

We present evidence from three contexts where professional networks have been developed at scale to improve teaching and learning. The networks have vertical relationships with different actors within schools and beyond.

In Kenya (Figure 3.1), Wasichana Wetu Wafaulu (Let Our Girls Succeed) is an FCDO-funded programme led by EDT. Teacher CoPs comprise a network of five schools that are facilitated by itinerant instructional coaches. The coaches decide the discussion focus, as informed by teacher priorities and lesson observations, to identify challenges and outstanding teaching to be shared.
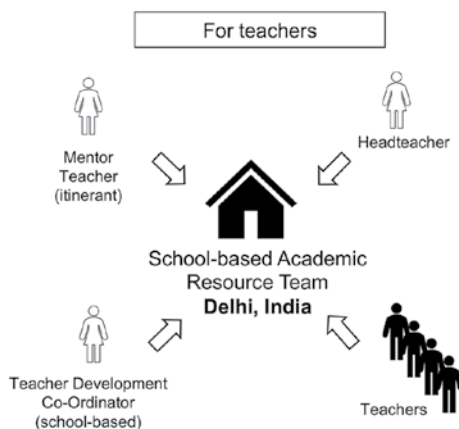
Source: Rossignoli et al. (2020).

Figure 3.1   *CoPs in the Kenya case study*

In Rwanda (Figure 3.2), the FCDO-funded Building Learning Foundations programme is led by the Rwanda Basic Education Board and EDT. Collaborative networks for headteachers and teachers are at the heart of the programme's philosophy. It uses a system leadership model, in which 480 high-performing headteachers – the Leaders of Learning – act as change agents to drive improvement across school systems. They

lead PLCs for headteachers each term across 15–16 schools. In turn, headteachers support termly school-based teacher CoPs, which are led by teachers and supported by school subject leaders (SSLs). These are highly structured and follow guidance in toolkits. Headteachers monitor CoP performance. The CoPs were rolled out over three phases, meaning that they were at different maturity levels during the research.
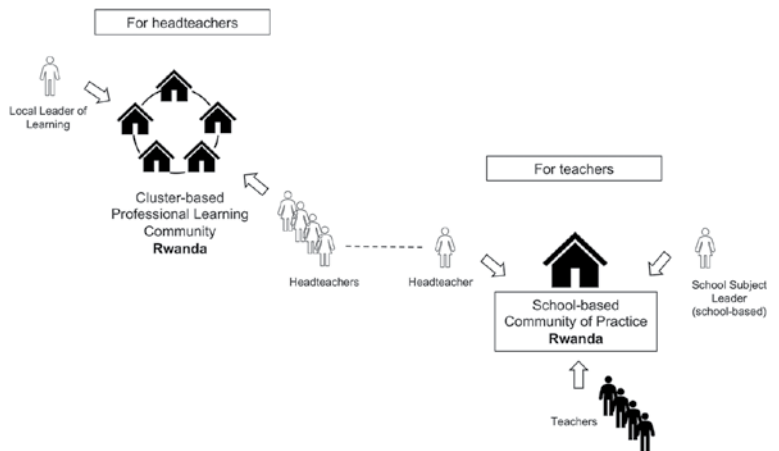
Source: Rossignoli et al. (2020).

Figure 3.2   PLCs and CoPs in the Rwanda case study

In India (Figure 3.3), STiR's teacher network model involves academic resource teams in each school, which are set up for teachers to share a professional dialogue on teaching and learning. A teacher development coordinator (TDC) in each school supports these; in turn, they are supported by mentor teachers (MTs), who are expert teachers who work across schools. The networks were explicitly designed to deliver pedagogical change at scale.

METHODOLOGY

The evidence presented in this chapter draws on four studies, shown in Table 3.2. Each of the studies examines how collaborative networks function and looks at the role of education leaders in supporting them. Note that the studies do not aim to track the impact of collaborative networks
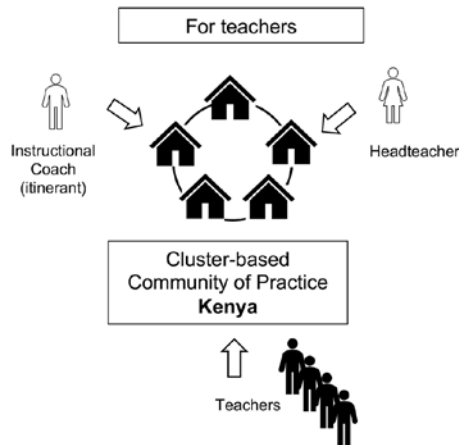
Source: Authors.

Figure 3.3   CoPs in the Delhi case study

on learning outcomes; however, all of the studies were commissioned for programmes which have independently verified evidence of impact on teaching and learning outcomes (DFID, 2020; Coffey International, 2017; Lanthorn, 2018).

Table 3.2   Empirical evidence from the four studies

Jurisdiction	Type of professional network	Study
India (Delhi)	Teacher CoPs, 1,000 schools	FCDO research, conducted by EDT
Kenya (five regions)	Teacher CoPs, 500 schools	Teachers Learning Together Research, conducted by EDT
Rwanda (national)	Teacher CoPs, 2,500+ schools	
	HTs Professional Learning Communities, 2,500+ schools	WISE research, conducted by Education Commission and EDT IIEP-UNESCO and EDT research

'Teachers Learning Together' (Rossignoli et al., 2020) (TLT) is a three-year study (2018–2021) of professional networks at scale in Kenya and Rwanda, investigating the characteristics of effective teacher

CoPs and their impact on teaching outcomes. It also looks at the role of mid-level leaders. In Kenya, the TLT study uses an experimental design that compares the treatment and control groups; in Rwanda, it employs a quasi-experimental design, comparing the variable CoP Impact for participants with different exposure levels from phased roll-out. The effects of CoP characteristics (as measured through teacher self-reporting) are assessed through a correlation analysis against CoP outcomes. The sample size is 610 teachers drawn from 17 schools across the two countries.

The WISE research (Al-Fadala et al., 2021) in Rwanda looked at the relationship between headteacher practices and equity during school closures and reopening in 2020, here with a focus on the role of headteacher networks. It used a survey of 100 headteachers and 21 district and sector officers on their leadership practices and school level case studies with 15 qualitative interviews of school leaders, teachers, students and caregivers and two central-level officials to explore school leader attitudes, practices, behaviours and beliefs as well as collaborative practices with other school leaders and with teachers and the relationship to continued learning and equity.

The IIEP-UNESCO and EDT research (forthcoming) looks at instructional leaders at the middle-tier of the systems, and their role in improving teaching and learning, here drawing on five case studies, including the Rwanda leadership of learning model. The study drew on ten in-depth qualitative interviews with leaders of learning and stakeholders to shed light on their practices and the characteristics of professional collaboration.

The FCDO research (Gibbs et al., 2019) looked at the process of scaling teacher networks in Delhi; it investigated the professional characteristics of mid-level leaders (TDCs and MTs) associated with effective scaling. Effective MTs and TDCs were identified by programme staff. Here, 60 in-depth qualitative interviews were conducted with TDCs, MTs, headteachers and government staff and then coded for the TDC and MT competencies associated with improved teaching and learning.

FINDINGS

On the Characteristics of Professional Collaborative Relationships

Trust and an open, collaborative culture

The TLT study found that effective CoPs[1] had a collaborative culture with equal participation opportunities and an atmosphere of trust; in Rwanda, a statistically significant positive correlation was found between an index of such conditions and teachers' self-reported CoP Impact (r=0.628 p<.001 n=97). The teachers described this open culture as follows:

> Cluster meetings motivate me because this is where I learn from other teachers [...] No one has exclusive rights to speak, and it's a free discussion. It motivates you to air your views. Also, you hear other people's views on methods of teaching. (Teacher, Kenya)

Similarly, the Delhi study found that openness within teacher relationships was a key shift resulting from the networks:

> We are very private in our classroom […] This is the main problematic thing for us like, 'Oh my god! We are going to expose ourselves, and they may point out some mistakes in us.' It is very important to remove that thing first. The mentors are doing it very well. They are making us relax. (TDC, Delhi)

Professional dialogue and a process of enquiry

Within this atmosphere of trust, constructive professional dialogue and reflection flourished. Across the studies, high-quality dialogue was identified as a key feature of effective collaboration and was characterised by a discussion of common problems, questioning, sharing reflections and feedback, seeking alternatives and planning of action points. In TLT, a statistically significant positive correlation was found between an index of such factors and self-reported CoP Impact (r=0.645, p<.05 n=278). The qualitative findings across the studies illustrate this further, showing how dialogue helps practitioners share ideas and solve problems:

> [We] talk freely. If you have a really burning challenge, you talk it out, and they help you. If there is one person who went through the challenge and suc-

---

[1] For baseline analysis, data from a teacher survey were used to create a CoP impact index by using principal component analysis. The index combines self-reported items on the impact of the CoP on teaching and student learning.

cessfully did the right thing they share and you will learn from them. (Teacher, Kenya, TLT)

We share our experiences and challenges. Through this collaborative style of learning, we will find more ways or solutions to situations or issues faced. (Leader of Learning, Rwanda, IIEP-UNESCO)

The WISE study captured similar dynamics during the COVID-19 crisis, finding 81% of leaders used professional networks such as PLCs to share advice with peers and that headteachers collaborated to find solutions for issues. In Delhi, the interviewees went further, describing faster feedback loops in solving problems:

The [network] is 12:40 to 1:00. In that time, whatever is discussed, feedback on that comes up on the very same day or in two to three days. This problem is there; it can be sorted out. Feedback takes less time to come up. We all then worked on it together. (Head of School, Delhi)

Across the studies, the professionals were highly motivated by this problem-solving dynamic:

We are very motivated […] We would find ourselves asking each other 'when is the CoP?' and when the CoP comes, we would go there running […] I'm very eager to share the difficulties encountered in a certain course I was teaching and for them to guide me on how I can resolve that. (Teacher, Rwanda, TLT)

Peer challenge and accountability
Professional challenges and peer accountability were part of this open relationship. The IIEP-UNESCO research shows how open critique in the Rwandan PLCs functioned as a way to support improvement, course correction and problem-solving:

We criticise everything that needs to be criticised. When we meet next time, some have made corrections, while others have nothing done; then, we offer more advice. This has helped many of our colleagues improve their schools by telling the truth to each other. (Leader of Learning, Rwanda)

Headteachers became more serious and started analysing their school performance data, and they discussed this with their teachers. All of us acknowledged that we have a role to play in the problem of repetition and dropout. We decided that we should correct ourselves and shall put it in our 'imihigo' [performance contracts] for this year. (Leader of Learning, Rwanda)

This sense of shared accountability came out strongly in the Delhi research, where a joint sense of responsibility to peers and for student learning was central to the relationship, especially when coupled with professional agency and autonomy:

> With the [teacher network], I feel accountable for supporting my colleagues. (TDC, Delhi)

> Our responsibilities have now increased to an extent, and we are changing our methods accordingly […] The responsibility of choosing the right practices has increased. (TDC, Delhi)

Shared purpose – a collective responsibility for improving student outcomes

Most of the studies find that the group's work – enquiry, challenge, dialogue, and so on – is driven by a moral purpose, a sense of collective responsibility for improving student outcomes:

> If [the students] pass out from the school and are excelling in the world […] then that could be my motivation […] I feel happy according to my conscience. (TDC, Delhi)

One MT eloquently described how it was clear intent and moral purpose that moved collaboration forward, even when the ways of working were ambiguous:

> Here, the basic thing is intent […] We are not doing it just for the sake of earning some money […] We are doing it because we want those children who are deprived; almost 90% of them are first-generation learners, so we want to do something for them […] It's not that we need everything in black and white because then, only we will move forward. The system is a little informal, and they know our intentions and we know their intentions. We can move, even in a phone call. (Mentor Teacher, Delhi)

The UNESCO-IIEP study describes the shift to shared goals, away from individual planning:

> Attending PLCs, I've learned the importance of analysing and planning together with teachers and stakeholders by sharing our goals and objectives. Because of this, in 2019, national performance was 100%, and we were in first place in government-aided school performance tables. (Headteacher, Rwanda)

The headteachers explained how a shared purpose, here focusing on the student, meant that 'real' issues were brought to the surface and complex problems solved:

> Before PLCs, the headteacher would sit alone and design, implement and monitor activities. Sometimes, these activities would not tackle the real issues and burden him alone. Now, parents are involved. We worked with parents and teachers to prioritise reducing the dropout rate of children from 8.4 to 3 in 2019 school years, which was a main challenge before 2019. (Headteacher, Rwanda)

The WISE study shows how shared purpose was critical during the Rwanda COVID-19 crisis, indicating that PLCs were a collaborative space to plan common goals to support school reopening: 70% of leaders said they collaborated with teachers to build a shared vision.

On Professional Collaboration and Leadership

All the studies examined the role of mid-level leaders in creating the conditions for successful collaboration.

> In the TLT study, support from midlevel leaders was correlated with effective CoPs. There was a moderate positive correlation between CoP Impact and an index of facilitation quality ($r=0.600$, $n=271$, $p < .001$ for Rwanda and $r=0.518$, $n=235$, $p < .001$ for Kenya). Similarly, headteacher engagement in CoPs is important: there is a statistically significant positive correlation between headteacher engagement and CoP Impact in both contexts (rho = .303 $p < .05$ $n=58$ in Kenya; rho = .233 $p < .05$ $n=97$ in Rwanda).

Across the studies, mid-level leaders supported professional collaboration in three main ways.

Offering expertise
Leaders offered expertise to support practitioners, such as through instructional guidance in Kenya, toolkits in Rwanda and role modelling in Delhi. The TLT shows that to get the benefit from CoPs, teachers needed some external stimulus or subject matter expertise, and this was provided by instructional coaches in Kenya and SSLs in Rwanda. The

teachers understood how important this external input was to avoid the recycling of poor practice:

> We have first-hand information other than getting information from another person; sometimes, the information can be distorted when it comes from one person to another. (Teacher, Kenya)

Facilitation

In all three settings, the leaders balanced this directive, standards-setting work with facilitative inputs. The TLT study investigated this balance. It looked at two leadership styles – directive and facilitative – based on the frequency of engagement with CoPs and in giving teacher feedback. It found that for SSLs offering subject expertise in Rwanda, more mature CoPs were positively correlated with directive inputs from these actors (rho=0.231 p<0.001), while less mature CoPs were negatively correlated (rho=-0.293 p<0.001). As CoPs mature, the importance of subject-specific leadership seems to grow. This suggests that once the CoP is established and teachers settle into routines, more explicit subject leadership could help with learning and classroom implementation.

The pattern is reversed for headteachers. For less mature CoPs, directive school leadership styles had a positive correlation with CoP Impact (rho=0.265 p<0.001); while for more mature CoPs, it was the opposite: directive leadership had a negative association (rho=-0.229, p<0.001). In other words, although headteacher input is instrumental, there is the danger of too much involvement as CoPs mature, which may obstruct their functioning. This may prevent teachers from developing a sense of ownership, who will see CoPs as an accountability mechanism rather than a learning community.

The studies offer rich insights into the shift leaders made in their styles, from traditional accountability figures to learning partners empowering practitioners to solve problems themselves. The IIEP-UNESCO research finds Rwandan leaders of learning were associated with profound changes, including promoting a collaborative learning culture across schools and districts, breaking down professional siloes and building collective responsibility:

> Before, every headteacher worked alone, but now, we share the experience. If I have best practices in my school, I share them with my colleagues. I became like a model; they came to see [...] how I have […] gained that practice. We have strong collaboration because we sit together, and we share what we can't do alone. (Leader of Learning, Rwanda)

> They used to say, 'I'm the headteacher and you are the teachers and you cannot tell me what to do.' But through the PLCs and through the local leaders of learning, they have now understood that they have to cooperate with the teachers and that it helps to be a team. (Sector Education Officer, Rwanda)

The Delhi study explores the TDC and MT characteristics associated with effective teacher networks and how they catalysed collaboration, pulling out the best from teachers and promoting high-impact collaboration:

> What I find is that the TDC himself or herself alone can't do anything. His basic responsibility is to create an environment where people should come out with ideas. (Head of School, Delhi)

The study finds high-performing leaders were differentiated by characteristics that include setting a clear purpose, facilitation and coaching skills, encouraging participation, openness to learning, ability to challenge the status quo and trust building.

Accountability

A more traditional accountability role continues to be an important leadership aspect that ensures CoPs flourish. In Rwanda, there was strong qualitative evidence that headteachers put in place accountability structures to ensure CoPs function and institutionalise learning, such as making time for CoPs and tracking the implementation of action points in class:

> [The headteacher] urg[es] us to apply what we have talked about in the CoP. (Teacher, Rwanda)

Similarly, the Delhi study found that headteachers' authority, as a sponsor for the teacher networks, was a critical success factor:

> [The headteacher] is with me so I can effectively [...] pass out my strategies to the teachers [...] Without the head of school, my impact will be 0%. (TDC Delhi)

DISCUSSION

Our findings describe the characteristics of successful professional collaboration in three settings. Features of lateral relationships between

Table 3.3   *How mid-level leaders enable professional networks*

Expert input	Collaborative facilitation	Accountability
Offer subject expertise	Set norms and expectations	Sponsor the network
Offer feedback and challenge	Support common vision	Ensure meetings happen
Bring external perspectives	Advocate for change	Follow up on actions from
Role model and shared best	Empower professional learning	networks
practice	Create trusting environment	Institutionalise continuous
	Foster inclusivity and	professional improvement
	participation	

Source: Authors.

peers are discussed, as well as features of the vertical relationships that professionals hold with mid-level leaders.

We find that when lateral relationships are characterised by four features – trust and openness, dialogue and enquiry, shared purpose and peer accountability – professional networks can create a learning and improvement environment, empowering practitioners to solve complex problems, such as how to improve teaching and learning. For example, the collaborations could deal with new challenges, as illustrated by head-teachers' work through PLCs during the COVID-19 response in Rwanda. The collaborative networks sparked professionals' motivation as a force for change and meant that professionals aligned their efforts towards clear outcomes, as the Delhi example illustrates.

In relation to vertical relationships, the dynamics are perhaps more complex. There is 'no one size fits all' for how mid-leaders – such as the TDCs, MTs, headteachers and instructional coaches – can promote effective collaboration. However we see three leadership functions emerging from our findings, as summarised in Table 3.3. First, in line with wider evidence, the leaders in our studies offered expert subject input to ensure the network did not recycle poor practice. Second, the leaders offered a strong set of facilitation skills to catalyse and structure effective professional collaboration. Finally, a more formal leadership and accountability role remained an important ingredient for successful collaboration. For example we saw the headteachers in Rwanda and Delhi, and the instructional coaches in Kenya, act as sponsor for the networks and take action to institutionalise professional learning within the school.

We have seen how it takes a matrix of different leadership roles – headteachers, subject experts, and so on – to offer the support outlined in Table 3.3. We saw that leaders need to shift their style and inputs accord-

ing to the context. For example, there is a danger in too much headteacher involvement in professional collaboration because it might obstruct the performance of more mature CoPs. We also saw how subject experts play a different role compared with headteachers and how their role may grow as CoPs mature.

Figure 3.4 proposes a simplified model of how change happens through professional collaboration, drawing on both the empirical evidence and wider literature presented in this chapter.
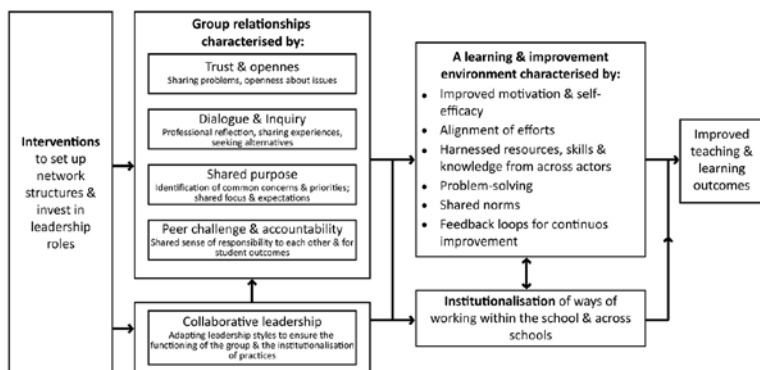
Source: Authors.

Figure 3.4   How change happens through collaborative professionalism: a simplified theory of change

CONCLUSION

The body of research synthesised in this chapter looks inside the 'black box' of delivery. It examines the relationships between different education actors – such as coaches, teachers and headteachers – both laterally between peers and vertically between practitioners and mid-level leaders. How can this evidence help us understand 'how change happens'? What are the implications for policymakers wanting to enable change at scale?

First, we argue that understanding relationships is critical to understanding how change happens within educational systems. The case studies show that when positive change is happening, relationships are characterised by more than an accountability-based dynamic. When peer relationships are characterised by four features – trust, professional dia-

logue, shared purpose and peer accountability – a learning and improvement dynamic is set in motion. This emphasis on the cultural and social elements of education relationships aligns with evidence of collaborative professionalism from higher-income contexts. Understanding these collective and informal elements helps us with ontological questions, broadening our understanding of what matters in education systems beyond a traditional emphasis on formal institutions, delivery structures and accountability relationships.

Second, we argue that the practitioner networks in the three case studies are more than a professional development intervention. As examples of collaborative professionalism, we see the relationship shifts as potentially transformative, representing system change. Actors were better able to deal with complex problems and align divergent views. Through the group dynamic, new norms and practices were established and institutionalised. Feedback loops were set in motion, where practitioners led cycles of improvement that made learning explicit and put new knowledge into practice.

Therefore, we see professional networks as vital tools for policymakers to bring about change at scale. But what practical action can policymakers take given that relationships within systems are constantly in flux and that positive dynamics are not guaranteed?

Our third argument is that practitioner networks need to be thought of as an open system; the support of actors from outside the network is critical for sparking positive change. This is an important entry point for policymakers wanting to catalyse change. In our case studies, these actors were mid-level leaders, such as headteachers, MTs and coaches, who played a key role in nurturing the group dynamics for improving teaching and learning. The case studies show the benefits of policymakers investing in distributed leadership structures, here understanding leadership as a matrix of actors – such as the MTs, leaders of learning, SSLs and instructional coaches – serving as change agents to catalyse effective collaboration. As the wider literature suggests, no single actor can offer the complex capacities required to shift learning outcomes.

Finally, we argue that these vertical relationships – between mid-level leaders and peer collaborations – are also based on more than an accountability dynamic. A collective spirit within mid-level leadership was important for achieving collective goals; rather than acting as an authority figure, leaders deployed facilitation and coaching skills to create a positive group dynamic and offer expertise in a constructive way. An accountability relationship was part of the leadership repertoire,

but leaders needed to understand when to deploy this according to the maturity of the collaboration.

These ideas offer fresh thinking for policymakers on how to strengthen delivery systems. The right policies can cultivate professional networks and change agents to spark positive disruptions. We argue that policymakers need to be intentional to build the capacities outlined in Table 3.3 and Figure 3.4. In other words, we urge policymakers to look beyond the more traditional focus on the 'hard-wiring' of education systems, such as structures and accountability systems. We call for further investment in the relationships and interdependencies between actors as outlined in this chapter, in other words the shared professional practices, processes and cultures that have the potential to transform teaching and learning.

REFERENCES

Al-Fadala, A., Amenya, D., Fitzpatrick R., Godwin, K., Kirby, J. & Korin, A. (2021). *Rwanda learning partnership: Insights on school and system leadership during COVID-19*. World Innovation Summit for Education (WISE).

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M. & Walton, M. (2017). From proof of concept to scalable policies: Challenges and solutions, with an application. *Journal of Economic Perspectives*, *31*(4), 73–102.

Barletta, B., Comes, D., Perkal, J., Shumaker, R., Wallenstein, J. & Yang, B. (2017). *Networks for school improvement: A review of the literature*. Columbia University Center for Public Research and Leadership, Columbia University.

Bodily, S., Chun, J., Ikemoto, G. & Stockly, S. (2004). *Challenges and potential of a collaborative approach to education reform*. RAND Corporation.

Bryk, A. & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. Russell Sage Foundation.

Bryk, A. S., Gomez, L. M., Grunow, A. & LeMahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better*. Harvard Education Press.

Checkland, P. & Poulter, J. (2006). *Learning for action: A short definitive account of soft systems methodology and its use for practitioners, teachers and students*. Wiley.

Coffey International. (2017). *Endline evaluation report step change window final report*. Evaluation Manager Girls' Education Challenge (GEC) Fund.

Cullen-Lester, K. & Yammarino, F. J. (2016). Collective and network approaches to leadership: Special issue introduction. *The Leadership Quarterly*, *27*, 173–180.

Darling-Hammond, L., Hyler, M. E. & Gardner, M. (2017). *Effective teacher professional development. Research brief*. Learning Policy Institute.

DFID. (2018). *DFID education policy 2018: Get children learning*. https://www .gov.uk/government/publications/dfid-education-policy-2018-get-children -learning

DFID. (2020). *Annual report, learning for all (LFA)*. DFID.

Donahoo, J. & Velasco, M. (2016). *The transformative power of collaborative enquiry: Realizing change in schools and classrooms*. Corwin and Learning Forward.

Donohoo, J., Hattie, J. & Eells, R. (2018). The power of collective efficacy. *Educational Leadership*, *75*(6), 40–44.

Dumont, H., Istance, D. & Benavides, F. (Eds). (2010). *Nature of learning: Using research to inspire practice*. OECD Publishing.

Eells, R. (2011). *Meta-analysis of the relationship between collective efficacy and student achievement*. Doctoral dissertation, Loyola University of Chicago.

Elmore, R. (2005). Accountable leadership. *The Educational Forum*, *69*, 134–142.

Elmore, R. (2008). Leadership as the Practice of Improvement. In *Improving School Leadership, Volume 2: Case Studies on System Leadership*, D. Hopkins, M. Nusche & B. Pont (Eds), (pp. 37–67). OECD Publishing.

Fullan, M. & Quinn, J. (2016). *Coherence: The right drivers in action for schools, districts and systems*. Corwin and Ontario Principals Council.

Gibbs, E., East, C., Hallgarten, J. & Jones, C. (2019). *System scaling in Delhi. A 'Learning partner' formative evaluation of STiR Education's Intrinsic Motivation model*. DFID.

Hargreaves, A. & Fullan, M. (2012). *Professional capital: Transforming the teaching in every school*. Teachers College Press.

Hargreaves, A. & O'Connor, M. T. (2018). *Leading collaborative professionalism*. Centre for Strategic Education Seminar Series Paper #274.

Harris, A. (2012). Distributed leadership: Implications for the role of the principal. *Journal of Management Development*, *31*(1), 7–17.

Hattie, J. (2016). Third Annual Visible Learning Conference: Mindframes and Maximizers. Washington, DC, 11 July 2016.

Hildreth, P. & Kimble, C. (Eds). (2004). *Knowledge networks: Innovation through communities of practice*. IGI Global.

Honig, D. (2020, December). *Bureaucratic barriers or administrative actions: The role of bureaucracies in successful education reform [Paper presentation]*. RISE Webinar. https://riseprogramme.org/events/bureaucratic-barriers -or-administrative-actions-role-bureaucracies-successful-education

Ianquinto, B., Ison, R. & Faggian, R. (2011). Creating communities of practice: Scoping purposeful design. *Journal of Knowledge Management*, *15*(1), 4–21.

Jensen, B., Sonnemann, J., Roberts-Hull, K. & Hunter, A. (2016). *Beyond PD: Teacher professional learning in high-performing systems. Teacher quality systems in top performing countries.* National Center on Education and the Economy.

Katz, S. & Earl, L. (2010). Learning about networked learning communities. *School Effectiveness and School Improvement*, *21*, 27–51.

Lanthorn, H. (2018). *Endline report on STIR communities of practices in Delhi and Uttar Pradesh*. IDInsight.

McAleavy, T., Thai Ha, T. & Fitzpatrick, R. (2018). *Promising practice: Government schools in Vietnam*. Education Development Trust.

OECD. (2017). *Systems approaches to public sector challenges: Working with Change*. OECD Publishing.

Priestley, M., Biesta, G. J. J. & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury Academic.

Pritchett, L. (2015). *Creating education systems coherent for learning outcomes*. RISE Working Paper.

Pritchett, L., Woolcock, M. & Andrews, M. (2010). Capability traps? The mechanisms of persistent implementation failure. *Centre for Global Development Working Papers*, No. 234.

Proger, A. R., Bhatt, M. P., Cirks, V. & Gurke, D. (2017). *Establishing and sustaining networked improvement communities: Lessons from Michigan and Minnesota. REL 2017-264*. US Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest.

Riley, K. & Stoll, L. (2004). Inside-out and outside-in: Why schools need to think about communities in new ways. *Education Review*, *18*(1), 34–41.

Rossignoli, S., Amenya, D., Jones, C., Kamana, D., Tiganescu, A. & Kudenko, I. (2020). *Teacher collaborative learning at scale: Governance and the role of school and system-level actors. Initial findings from the Teachers Learning Together study in Kenya and Rwanda. July 2020*. EDT.

Senge, P. (1990). *The fifth discipline: The art and practice of the learning organization*. Currency.

Silvia, C. (2011). Collaborative governance concepts for successful network leadership. *State and Local Government Review*, *43*(1), 66–71.

Spillane, J. P. (2006). *Distributed leadership*. Jossey-Bass.

Stoll, L. & Temperley, J. (2009). Creative leadership: A challenge of our times. *School Leadership and Management*, *29*(1), 65–78.

Williams, M. J., Leaver, C., Mansoor, Z., Qarout, D., Bilous, A., Mundy, K., Asim, M. & Bell, S. (2020). *Delivery approaches to improving policy implementation: A conceptual framework*. Education Commission, University of Oxford, UKAid.

Wohlstetter, P., Houston, D. M. & Buck, B. (2014). Networks in New York City: Implementing the Common Core. *Educational Policy*, *29*(1), 85–110.

4. Global education funders' perspectives on the potential of systems thinking to change education practices and achieve mass learning gains

Laura Savage, Clio Dintilhac, Raphaelle Martinez, Tjip Walker and Jason Weaver

INTRODUCTION

Children are in school, but not learning the basics.

More funding is needed to tackle 'the learning crisis', but trends show that more money does not lead to more learning.

If we keep going as we are, then there's no hope of meeting SDG4.

This was the fairly depressing narrative circulated in the boardrooms and inboxes of global education aid architecture back in 2016. After the high of agreeing on ambitious global education goals at the World Education Summit in Incheon in 2015 (a more sector-specific roadmap to the global 2030 Sustainable Development Goals), focus turned to the task of achieving these goals. It was becoming clear that something would need to be done differently, but there was no consensus on what. A small group of people representing different organisations across the global education aid architecture started to explore systems thinking as a way to build understanding at the country level of why learning levels were so low.

Some of these were people already attracted to systems thinking and who saw system diagnostics as a tangible entry point for introducing this thinking into the education sector. Others were interested in trying

something new. Over the course of the following three years, this group met irregularly in the margins of other global education conferences, each time reflecting on a new think-piece, report or process that their organisation had been involved with.[1] In April 2021, representatives of four of the organisations the most involved in this process spoke at the annual conference of the Comparative International Education Society and reflected on the value of systems thinking in their work. This chapter summarises what they said during that panel.[2] Thus, this chapter offers a curated perspective of global education aid actors who have self-selected as being interested in systems thinking. It highlights some of their incentives, challenges and reflections on the value of thinking about education as a system.

BACKGROUND

These global aid actors start from an instrumentalist view of education systems: the one major purpose of these systems is seen as enabling children to learn so that they can go on to lead productive lives, with all the associated correlations for other markers of health, rights and prosperity. They also recognise that every actor within a particular system – be that a national education sector, a classroom, a network of teachers or the global aid architecture – will have different and likely plural definitions of what a successful outcome from education is (i.e., it may not be 'learning outcomes' for all or at least the same definition of a learning outcome). It is also important to be clear that their perspective is one of 'outside in': they are a part of the global education system. When they talk about supporting change in national education systems, they can identify a lever or nudge movement within a system.

On a scale of expert to sceptic, most of these actors have put themselves somewhere left of the middle and self-identify as 'enthusiasts' regarding systems thinking. They have been trained in various disciplines: economics, political science, ethnography and history. As they explain, they came to 'systems thinking' as literature at various points

---

[1]   These included a debate panel at the 2018 Building Evidence in Education workshop in Oxford, UK, a side event during RISE Conference week in June 2019, and a small group meeting at the Gates Foundation during Education World Forum week in January 2019.

[2]   Clio was not able to attend the panel but contributed to this chapter subsequently.

in their careers. Systems thinking has grown as a stream of the literature specifically applied to international development, particularly by those working on evaluation or those who want to encourage the aid industry to embrace and work with complexity (Mosse et al., 1998; Green, 2016; Ramalingam, 2013). Yet in education, there has been a growing sense of frustration at the slow pace of outcomes (this despite the incredible success at getting children into school in the twentieth century).

Research on the failed scale-up of a successful contract teacher project in Kenya has been a major reference point for years (Bold et al., 2018). A team of researchers tracked a project that funded 'contract' (i.e., fixed-term hire) teachers in rural Kenya, a project first run at a small scale by an NGO and then scaled up by the government. This was an important and useful study and one of the most talked about in global education policy discourse for years. However, the researchers' findings here did not seem surprising to this group of actors (Elmore, 2016). Their varied experiences in working to support national education (and other public policy) reform had taught them that a carefully controlled intervention designed to respond to an identified problem and run by an organised NGO would yield results in a way that running it on a large scale through government systems would not. They could reflect on why this was the case as much as the economists running the study could: implementation fidelity, competing and contradictory demands on government officials, and different strains and dynamics within large-scale systems. However, these global actors on this panel were left with the following question upon reading this research paper: (how?) can (we?) change the system so that education outcomes are improved and for a mass of children and in a sustained way – that is, that the improvement repeats not just for one cohort but in an enduring way for that targeted cohort and again for cohorts behind them? 'Systems research' became the bucket term for research that would help to understand these 'how' and 'why' questions of sustainability and scale.

The rest of this chapter is presented in an interview format. First, we will introduce the 'speakers'. Clio Dintilhac is a senior programme officer at the Bill and Melinda Gates Foundation who focuses on the foundation's learning data portfolio and governance investments in education in the developing world. Raphaelle Martinez works at the Secretariat of the Global Partnership for Education; she currently leads the Education Policy and Learning Team, whose mandate is to mobilise and support the use of evidence and good practice in the fields of system strengthening, teaching and learning and equity and inclusion. At the time of this con-

ference session, Laura Savage was the Senior Education Adviser at the UK Foreign, Commonwealth and Development Office and was given a technical leadership role on 'education systems'. Tjip Walker is a senior policy analyst at USAID. Jason Weaver is a Senior Economist at the World Bank's Global Education Practice and leads its new Accelerator Initiative, which draws on the approach to education systems set out in the 2018 World Development Report on Education. The discussion was curated by Moira Faul. Panellist views are all offered here as personal reflections and do not represent the official policy of their organisations. However, they are all also recognised as careful thinkers on this topic, both within their organisations and across the global education aid space.

What Originally Attracted You to Systems Thinking? What about it Appeals to You?

Tjip: My role is to support teams across USAID to see the value of systems thinking in their day jobs. I will give an early example of when a colleague and I worked with a country team to offer them analysis of a particular sector's complex problem. The country team felt they understood the country enormously well and, therefore, would not be needing this kind of additional analysis to make their own decisions. They said to us, 'Go ahead, but we do not expect much out of you'. So recognising that we did not have an awful lot to lose, my colleague and I applied some systems models. We did the analysis and presented it to the team, including the ambassador. Much to our pleasant surprise, we identified some new factors and incorporated dynamics that they were unfamiliar with. We showed where there were tipping points and, at the end, they were stunned by our taking this thinking and applying it to something they thought they understood really well. This was my realisation that this was not simply something to read about in books but actually something that could be applied in the development context.

Laura: I was working in contexts where the word 'reform' was not palatable, where parents pushed back on evidence-based pedagogy because they wanted rote learning for their children and where private sector actors and donors were pulling in different directions. I was supporting the development of education sector plans in extremely different contexts but was finding that these somehow looked the same: with three ubiquitous goals of 'access, equity and quality'. Over my career, it has become more and more of a mantra of mine that we need to work differently in education to achieve learning gains. Working as an adviser in country

on large-scale reforms for a ministry of education and for different aid agencies, I was looking for a way to think harder about how and why good ideas, good intentions and money were not translating into strong education processes, structures and outcomes. I was attracted to systems thinking intuitively, being trained as a historian and political ethnographer. Systems thinking was attractive more for what it did *not* mean, which aligned with my experience – that there is not one shared outcome desired by all actors in an education system, that there is no such thing as a 'solution' and that a system is not some static entity that can be 'reformed' by outside-in or top-down reform.

Clio: The Gates Foundation is a new and very modest actor in education outside of the United States. The work on education in sub-Saharan Africa and South Asia started in 2018 and is a fraction of what the foundation spends on US education. As a novel actor, we were keen to better understand the gaps facing education systems in the developing world to improve their reading and math outcomes in primary schools. From our work in health in the developing world, we had a sense that we were unlikely to find 'silver bullets' to improve service delivery at scale in the public sector. We also realised that there was not really one authoritative framework in education, for example, the equivalent of Financial Sector Assessment Programme in the financial sector that informed policy. This may be the case for many reasons, for example the systemic nature of the financial sector and the types of risks associated with a financial crisis are distinct. The decision was for the programme to start by focusing on supporting the research and tools that would help inform and prioritise the government's education policies. As part of this first wave of investment, we supported the World Bank's Global Education Policy Dashboard and the Research on Improving Systems of Education (RISE) programme, jointly with the Foreign, Commonwealth and Development Office (FCDO), two endeavours that were trying to think at the system level. Since then, the programme has evolved to invest more at the country level, including the World Bank's Accelerator imitative and, in general, advocacy around the learning crisis. At a more personal level, I had experienced first-hand, in places like Ghana and Niger, the fragmentation of policymaking. Therefore, I was particularly sensitive to the notion that a system needs to be clear on the goal it is trying to achieve in priority and on the need for the different parts to work coherently and through iterative adaptation towards that goal.

How has the Word 'System' Landed in Global Education Discourse and Practice in Your Organisations?

Laura: I do not use it anymore. In 2014–2015, I tried to use the language of 'a systems approach'. However, I got a spectrum of responses, from 'I don't like to think about systems stuff because it is too messy; don't talk to me about complexity' to 'but we don't need a new approach, because we work at the system level already: we do budget support'. However, 'system' does not mean 'national', 'big' or 'sector'. I find that I get more traction on the underlying meaning of 'system' if I use the language of 'thinking and working politically' within my organisation or if I use 'test, learn, adapt' in country. In the 2018 then-DFID Education Policy, we applied these ideas to education in a four-pronged way. An education system first requires thinking about the inputs. Sometimes, systems thinking in education is made out to be an argument that we do not need more money in global education; that we need to do more with what we have. However, we do need more funding and more inputs. We then need to think about processes: the building blocks in state systems that make these inputs effective or not, such as data or procurement or public financial management processes. Third, we need to think about the people in the system: the relationships between the education stake-holders, which range from accountability relationships (such as what the RISE systems framework focuses on) to professional collaboration to personal. The fourth layer is politics, which requires recognising that there is a whole swirl of 'big P' and 'little p' politics affecting implementation and – crucially – that this needs to be understood but cannot necessarily be changed. The first layer is the easiest, but we are not going to change anything unless we tackle all four. So we found a way to approach systems thinking without using the 'S' word.

Jason: I am fond of analogies, so I will offer one here. We are all familiar with the conceptual definition of a nation – a people united by shared history, lineage, culture or language and who are inhabiting a particular geographic area – but the concept of a nation becomes infinitely clearer when you see a map. At the World Bank, our 2018 World Development Report (World Bank, 2018) – Learning to Realise Education's Promise – put forth some of our clearest thoughts on how to make systems work for learning. Though it does not push the 'systems' language or thinking explicitly, the WDR 2018 was able to capture all of the challenges and complexities that those of us on this panel know deeply, but it did so through a relatively simple and concise framework that was notably

digestible to actors outside or new to the education sector (e.g., Ministries of Finance). It posits that struggling education systems are typically missing one or more of the four learning ingredients at the school level: prepared learners, quality teaching, inputs and infrastructure and school management. It also put forth three relatively simple recommendations that also reflect a systems approach: (1) we must assess learning, which should be a serious goal of any education system, to understand where we are at and where we are headed; (2) we must act on evidence to ensure schools work well for learners; and (3) we must align actors – which is far more complicated than one might think – in a common and relentless pursuit of student learning.

In the wake of the 2018 WDR, we still needed a 'map' to help better clarify and apply the conceptual framework to actual education systems. The Global Education Policy Dashboard is our map. It applies a three-dimensional approach by looking at practices (or service delivery), policies and politics and then identifies a set of indicators that are focused on the same four school-level ingredients: teaching, school management, inputs and infrastructure and learner preparation. The Global Education Policy Dashboard offers a unique compilation of indicators that is comprehensive yet still focused enough to help stakeholders pay attention to what matters the most for learning.

Raphaelle: The word 'system' alone is not getting a lot of traction. If we were to say 'systems approach' or 'system thinking' to our country partners, they would ask the following: 'Ok, how is that different from what we have been doing before?' 'How is a system different from a sector?' It remains theoretical, and these distinctions between sector and system do not lead us very far. It is also not clear what a systems approach would mean in terms of analytics, implementation and operations that would be different from the more traditional and rather standardised way the international community has supported educational planning, implementation and monitoring for the past few decades. However, we hear a demand for big changes: more than a year ago, in the context of developing our new Global Partnership for Education (GPE) strategy, we held a roundtable at which the Ministers of Education of our country partners said, 'We have to be bold; we have to be transformative'. So for us, it is not about defining what a system is, but rather, it is reflecting and leveraging the assets of our partners to think and work differently to achieve system transformation. This is a way of thinking that looks at unpacking complex issues, the interdependencies of various system components and

their alignment and that aims to identify the levers and opportunities to unlock system-wide changes.

For me, as an educational planner by background, it is about enabling and incentivising a more agile, quick, responsive, targeted way of doing policymaking in which there is a premium on integrating effective delivery; identifying potential ripple effects of the policy and programmatic priorities education stakeholders align behind; aligning actors' resources of different kinds: leadership, finance, expertise and knowledge. For this, we need data, evidence (1) on the pockets of sector underperformance, (2) understanding the causes of the sector issues and bottlenecks, and (3) who is doing what and whether it is working so as to build strong accountabilities within the system. It is also about stepping back slightly from a comprehensive sector-wide approach to focus on a few key priorities to lay down a reform agenda because nowhere can we do everything at the same time. But then, how do we build consensus around priority reforms? How do we accommodate the different perspectives and behaviours of all the actors involved because there will not be full convergence of interests and resources? How do we understand the root causes of weak education outcomes to be able to come up with that prioritisation agenda? How do you implement these priorities – through testing and trial and error and nudging and working out leverage points? Our new strategy is about trying to put the word 'system' into practice.

Tjip: Education is by no means the only sector where there are issues. I have a vantage point across different sectors. In USAID, 'system' is not a word we shy away from. It is a matter of policy to strengthen what we call 'local systems', which can be defined depending on what the problem is. We have a theory of how we get to sustainable change: it is about strengthening the systems that are capable of being able to produce results in a way that allows local actors to perpetuate and realise particular outcomes. The market system is a well understood and certainly mainstream concept for agricultural economists and those who work on food security. However, in the health sector, we still have challenges between those working on improving health systems and those who focus more on trying to eradicate particular diseases – and these sometimes clash, and there are trade-offs. However, the important thing is that systems thinking is actually a set of related concepts, so you do not have to focus on the term 'systems' if you are essentially applying systems practice.

I would make two observations about my experience in education. First, there is a tendency to look out for 'solutions' that can be introduced into a system to achieve better learning outcomes. From a systems point

of view, you need to start by understanding that system in a particular context and the opportunity for change, who is motivated to make change and – something we often neglect – who holds the power to make change. For outsiders trying to promote change, the challenge is unlocking the already pent-up desire for change. Diagnostics are useful for this because they give you a set of tools to understand where that opportunity is. Or this may not be the case: it may be that there is simply not going to be the opening for the kind of change you are talking about. Second, I see a lot of effort looking at the impact of particular interventions and predicting the results from 'what works'. The concept of 'emergence' in systems thinking is that in complex systems, we have very limited ability to make predictions about how the system is going to operate, and understanding causality in a complex system is only possible to a small extent. This makes policymakers and funders a little nervous. The comments earlier about the importance of adaptation, learning and leverage points are all recognising the fact that contexts really vary and that it is important to understand dynamics and have the capability to respond. So it is not about whether we adopt a word or not, but it is also whether or not you are willing to adopt a set of concepts and behaviours that go along with some humility about our ability to claim what we can accomplish, even with large amounts of money.

Clio: The Gates Foundation has been investing in improving health systems for some time. In health, there was this impression that the task of funds like the Global Financing Facility, which aims at improving health systems, was much more complex than the ones of 'vertical' funds (like the Gavi, Global Fund) that focus on specific diseases or on the delivery of vaccines. In the global education work, the Gates Foundation decided to focus similarly on a narrow section of the education system, that is, improving reading and math in primary schools, which was as close to a 'vertical' fund that we could get: in a way, improving reading and math outcomes may be the part of education for which we have a solution that is closest to a vaccine, that is, structured pedagogy with regular monitoring. So our focus has really not been to improve the education system as a whole but really just that narrow sliver. The part of the work that went deeper into system change has been in the context of our investment in the Central Square Foundation (CSF) in India. In that context, CSF designed a diagnostic tool, very much like 'root cause' analysis that examined how different components – for example, pedagogy, academic support and expectations – contributed to low learning outcomes. This type of very practical, voluntarily very simple tool was

one example of an interesting approach to engaging government stake-
holders around a diagnostic of their system.

**So How Can Global Education Funders, Mostly in the Global
North, Handle the Tension Between Wanting to Strengthen
Education Systems in the Global South and Being Outsiders to
those Systems?**

Jason: To be clear, everyone wants the desire to change to come from
within – from governments, NGOs, the general population, and so forth.
However, we also know that every politician, certainly every minister of
education, will say they want education outcomes to be better, but are
they willing to push through the politically difficult reforms that may be
needed? Strong political commitment at and above the level of the minis-
try of education to improving foundational learning is the sine qua non of
the Accelerator Initiative. We know it is difficult to quantify or measure,
but we must be honest about the expected pace of progress. We look at
ways that we can help use data, as Tjip mentioned, to better understand
the opportunities for locally led change. Whether through international
benchmarking (e.g., Learning Poverty as a metric, the Human Capital
Index, PISA scores, etc.) or by starkly showing what kids are and are not
capable of doing in terms of age-appropriate skills (e.g., the citizen-led
assessments like the Annual Status of Education reports in India), this
can be an eye-opener and help galvanise political will. In practical terms,
many of the necessary reforms are deeply political. Curriculum reform
can take several years to unfold; parental perspectives and social norms
can be enablers or blockers. There are issues that sit outside of what we
think of as technical education changes but are all factors that matter.
We also need actors outside the government to establish accountability
and sustain commitment across political cycles. This is what a 'systems
approach' tries to recognise.
Laura: Moira [the panel chair] wrote a report for us [DFID] a few years
ago, and it was based on a series of interviews with global and national
education actors. It was striking that those who work in country, who
grapple with issues of implementation politics and governance every
day, were much more able and willing to 'think' systems. The further
out you get from 'the field', maybe it becomes easier to think in linear
ways because at the global level, the 'system' is just so massive. Some of
the writing on education systems that I have enjoyed most has been that
which takes the classroom as a system, as a set of interactions between

teachers, community members and students. All those layers I talked about earlier – the inputs, processes, people and politics – are all there. So I think it is less about 'outsider/insider' when thinking about an education system and more about framing the system and trying to identify what the RISE programme would call 'coherence' within that system. It is about understanding who are the decision makers and those who hold the potential to change and sway some of that system: this is key for anyone (inside or outside) who is trying to affect change within that system. And it will change constantly.

Raphaelle: As a partnership, GPE brings together donors, multilateral institutions, civil society, teacher representatives, philanthropic foundations and the private sector behind partner country governments' policy priorities. We promote evidence-based policy dialogue at the country level to identify the key education priorities that have the potential to unlock system-wide change. Our operating model is also designed to be country driven, context sensitive and raise awareness around political economy questions. In essence, what we do is work through our partners to support the use of relevant evidence and knowledge, to provide technical support and to build the capacities for inclusive and informed in-country policy dialogue and policymaking.

We also build incentives into our funding for actors to ask themselves the right questions and stimulate what we call the 'enabling factors' to system transformation. The enabling factors are areas that are key in shaping sound foundations for effective education systems: data and evidence, planning and monitoring, sector coordination and alignment and domestic finance. Through these enabling factors, we want actors to drive a contextualised dialogue around the critical challenges across the enabling factors that would limit their effectiveness in supporting particular policy priorities. These policy priorities are up to the governments. We have only broad policy priorities at GPE – for example, 'learning, equity, inclusion' – and this breadth is on purpose because it puts a marker where we want to see progress, but it is up to the country partners to work out the specifics of what and how.

Tjip: Understanding perceptions and perspectives is important, and this has to be one of the differences that defines a systems approach, which recognises that all actors in a system have views about how that system is supposed to perform. If our engagements are only at the ministerial level, then we miss important pieces of the puzzle. However, I would also underline how difficult change in complex systems is. If we are working within a complex system, we must recognise the number of

factors at play and that there are forces pushing for progress and forces pushing against progress. As a principle of systems thinking, this means that we need to have some humility about what is actually possible: in some places, there may be more orientation than change than in others. One important construct that comes from systems thinking is called 'the iceberg model'. This explains that below the way we construct systems are deep understandings about how we want society to operate, what our aspirations are and how we think change works. We carry all of these mental models around. I think one reason that systems thinking has perhaps had some difficulty making inroads into the international education sphere is because of some of the mental models that many people hold about the nature of change. It is difficult to get rid of those mental models. However, it is important to understand what these are. Another benefit of systems thinking regarding this insider/outsider question is that it puts a real focus on the nature of relationships between actors. Understanding these – and going beyond accountability, as the RISE framework focuses on – is an important consideration.

Clio: I think a discussion we had some time ago in this group was about whether global funders could be helpful if they associated their funding with the right 'diagnostic tools or approaches'. These diagnostics would enable us to target aid money in a more efficient way and inform the country's policies. Education sector analysis, which was a key tool in the last GPE operating model, is an example of this. However, I think there was also a recognition that if you associate the diagnostic with money, there is a risk it will be done in a perfunctory way and not owned. I think another risk was that the diagnostic tools ended up being academic tools trying to show the complexity of the system rather than being easily understandable prioritisation tools.

To better understand how improvement had happened at scale in country and what type of useful support external donors provided, we supported the Learning@scale study, which looked at the minimal components needed to see the impact of donor-funded projects and the system enablers for scale (Stern et al., 2022). We also supported Luis Crouch in putting together a review of how government-led programmes in the developing world, in particular Sobral in Brazil, Puebla in Mexico and nationwide in Kenya, had improved reading outcomes in their public education system (this was published by the RISE programme, see Crouch, 2020). That study has several take-aways, but I will mention three. First, consistent with what Jason is saying, there is the fact that external actors/ donors certainly can come as a form of support, as in the case of Kenya,

but there is a need for strong internal motivation to enable effective reading reforms. Second, none of these systems provided a formal diagnostic of the system. Their diagnostic was intuitive and implicit, closer to Laura's framing of 'thinking politically'. Finally, improvement was not achieved by rehauling the entire education system. These exemplars instead thought tightly about what needed to change in the classroom and then went on to address the couple of system enablers needed to make things work. There is possibly a role for donors in sharing these lessons and putting what happens in the classroom back at the centre of the discussions. These take-aways have informed our own investments; for example, we recently supported the research organisation Research Triangle Institute International (2018) to put together (in a project called 'Science of Teaching') all the evidence we have on structured pedagogy for the benefits of practitioners and policymakers.

How Does Systems Thinking Help Us Think About and Achieve Sustainable Change?

Tjip: One of the tools we used at USAID to help our systems thinking is a relatively simple construct – the 'five Rs'. This provides broad categories for trying to understand systems. One R is 'results' – the results you are trying to seek. Another is 'resources' – those that go into the system on a continuing basis as necessary, such as financial and human resources. There needs to be a positive translation here: people need to feel positive about the results so that they will be supportive of more resources going back into it. So there is a feedback loop. These resources are essential for sustainability. Donors can provide these on a temporary basis. A third 'R' is 'rules' – the rules that govern the system. It might be easy to change the rules, but what really matters is how they are enforced, and this is fundamental in determining whether change is sustainable over time.

Laura: From a slightly different perspective, I hear this question a lot from Ministers of Education and Ministers of Finance, as well as from donors: we have put funding into education, but we are not seeing learning levels increase. Some of the drive for a different way of thinking about education comes from this desire for a sustainable impact of investments already made. Over the past year, the World Bank and FCDO have partnered together to support the Global Education Evidence Advisory Panel in publishing a report on 'Smart Buys' in global education. This summarises the small evidence base with cost data to help highlight some

ideas that might work, if tested/adapted, and those that might not because we have lots of evidence that they do not. There is an interpretation of that report that says, 'Here is an intervention, go do it everywhere, and it will work', but there is a carefully worded series of paragraphs that explains these ideas need to be built into context, adapted and tracked – in other words, implemented through a 'systems approach'. What I think we have less evidence of and what I think matters most for the question of sustainable change is why: why did an intervention/reform/idea work in practice or not? This kind of research would capture the coincidences, the accidental stars aligning, the feedback loops and the champions for change that make something happen. This is the sort of information that helps to explain how a system works, not so that we can replicate it, but so that we can understand why learning has improved (or not).

CONCLUDING REMARKS

This chapter is just a snapshot of a conversation that a group of education aid actors have been having for nearly a decade and will no doubt continue to have. They take a very practical perspective on systems thinking, and in this spirit they offer some tools that they found useful in their work. Not all of these are framed explicitly as 'systems thinking' tools, but for each tool, authors here have found them useful to get across the key concepts that they find compelling.

- USAID's '5Rs' framework for systems practice is a practical tool to promote good systems practice. It highlights the five key dimensions of systems: results, roles, relationships, rules, and resources. Collectively, these 5Rs can serve as a lens for assessing local systems and as a guide for identifying and monitoring the interventions designed to strengthen them (USAID, 2016).
- The World Bank's World Development Report 2018 – *LEARNING to realise education's promise* – is its first devoted entirely to education. The report leverages systems thinking in various ways and argues, inter alia, that countries must address stubborn system-level technical and political barriers for new reforms to sustainably improve learning (World Bank, 2018).
- The Global Education Policy Dashboard: Although this is not a dynamic system diagnostic tool, it is meant to bring together key datapoints to inform the diagnostic of an education system (World Bank, n.d.).

- The 2018 DFID Education Policy: This, as Laura outlined above, is a simple four-layered framework to think about how a planned intervention will fit within the dynamics and available levers of a particular system (DFID, 2018).
- The new GPE operating model 2025 sets out a 'system transformation' approach that is based on an emerging body of evidence on how to catalyse systems-level change, including some of the references cited above. GPE supports partner countries in transforming education systems at scale to get the most vulnerable girls and boys in school, improve teaching and learning and build education systems capable of withstanding shocks and adapting to new challenges (GPE, 2020).
- 'System diagnostic' tools have proliferated in recent years (see Chapter 10, this volume, for a deep dive into one). Fundamentally, a 'system diagnostic' should help a community that is grappling with a complex problem to come to a common understanding of why it might be happening, hence giving them the ability to address the real problem. There are several tools in this space – from the building state capability approach to problem-driven-iterative adaptation (Harvard University, 2022) to Oxfam's practical guidance to systems thinking (Bowman et al., 2015) to some education-specific adaptations (such as UNESCO IBE's framework (UNESCO, 2012), chapters in UNESCO/UNICEF's (2021) Education Sector Analytical Guidelines on institutional and political analysis and a tool based on the RISE framework (RISE, n.d.); for more see Chapter 10, this volume). With system diagnostics, the process is can be even more important than the tool.

REFERENCES

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A. & Sandefur, J. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics*, *168*, 1–20. 10.1016/j.jpubeco.2018.08.007.

Bowman, K., Chettleborough, J., Jeans, H., Rowlands, J. & Whitehead, J. (2015). *Systems thinking. An introduction for Oxfam programme staff*. Oxfam. https://oxfamilibrary.openrepository.com/bitstream/handle/10546/579896/ml-systems-thinking-151020-en.pdf;jsessionid=28872BACEC5FD736EB387C4A2442F273?sequence=1

Crouch, L. (2020). *Systems implications for core instructional support lessons from Sobral (Brazil), Puebla (Mexico), and Kenya*. RISE Insight Series. 2020/020. https://riseprogramme.org/sites/default/files/publications/20200708_RISE%20Insight%202020_20_Crouch_Systems.pdf

Department for International Development. (2018). *DFID education policy: Get children learning.* https://www.gov.uk/government/publications/dfid-education-policy-2018-get-children-learning

Elmore, R. F. (2016). 'Getting to scale …' it seemed like a good idea at the time. *Journal of Educational Change*, 17, 529–537.

Global Partnership for Education. (2020). *GPE 2025 operating model framework.* https://www.globalpartnership.org/sites/default/files/document/file/2020-11-GPE-Board-meeting-operational-model.pdf

Green, D. (2016). *How change happens* (p. 288). Oxford University Press.

Harvard University. (2022). *PDIA toolkit.* https://bsc.cid.harvard.edu/PDIAtoolkit#:~:text=Problem%20Driven%20Iterative%20Adaptation%20(PDIA,adapt%20and%20then%20act%20again

Mosse, D., Farrington, J. and Rew, A. (Eds). (1998). *Development as process: Concepts and methods for working with complexity*. Routledge.

Ramalingam, B. (2013). *Aid on the edge of chaos: Rethinking international cooperation in a complex world.* Oxford University Press.

Research Triangle Institute International. (2018). *How to guides.* https://scienceofteaching.site/how-to-guides/

RISE. (n.d.). *The RISE system diagnostic.* https://riseprogramme.org/tools/rise-system-diagnostic

Stern, J., Jukes, M., Piper, B., DeStefano, J., Mejia, J., Dubeck, P., Carrol, B., Jordan, R., Gatuyu, C., Nduku, T., Punjabi, M., Harris Van Keuren, C. & Tufai, F. (2022). *Learning at scale: Interim report*. RTI. https://shared.rti.org/content/learning-scale-interim-report#modal-72-2903

UNESCO. (2012). *General education quality analysis/diagnostic framework (GEQAF).* http://www.ibe.unesco.org/sites/default/files/resources/geqaf-2012_eng.pdf

UNESCO, UNICEF, Global Partnership for Education, & Foreign Commonwealth & Development Office. (2021). *Education sector analysis. Methodological guidelines.* http://www.iiep.unesco.org/fr/publication/education-sector-analysis-methodological-guidelines-vol-3-thematic-analyses

USAID. (2016). *The 5Rs framework in the program cycle.* https://usaidlearninglab.org/sites/default/files/resource/files/5rs_techncial_note_ver_2_1_final.pdf

World Bank. (2018). *LEARNING to realize education's promise.* https://www.worldbank.org/en/publication/wdr2018#:~:text=at%20the%20center.-,The%202018%20WDR%20explores%20four%20main%20themes%3A%201)%20education's%20promise,make%20systems%20work%20for%20learning

World Bank. (n.d.). Education Global Practice Department. *Global Education Policy Dashboard.* https://www.educationpolicydashboard.org/sites/epd/files/resources-documents/GEPD%20Implementation%20Brief.pdf

PART II

From theories to systems thinking and
practices

5. Understanding travelling reforms from a systems perspective

Gita Steiner-Khamsi

Arguably, the malleable use of the term 'systems' has been a cause for great confusion and unnecessary controversies among international cooperation experts. It might therefore be helpful to start out this chapter by pointing out what the 'systems approach' does not entail, followed by a brief sketch of key system-theoretical notions used in Niklas Luhmann's sociological systems theory. The bulk of this chapter is devoted to explaining four key concepts or set of concepts in comparative policy studies: externalisation, projection, reference and counter-reference societies, and reception and translation. All these terms have been inspired by, but do not strictly follow, sociological systems theory. In other words, I do confess without any remorse to my creative handling of sociological systems theory.

THE OTHER SYSTEM APPROACHES

As mentioned above, some boundary work is necessary to distance oneself from how the term 'system' is commonly appropriated, misinterpreted and at times criticised in the field of international and comparative education. In the field of applied comparative policy studies, the term 'system' is often used interchangeably with 'national education system', and by implication with policies that regulate that 'system'. Most visibly, the World Bank's SABER (Systems Approach for Better Education Results) project provokes that association with its name. In effect, SABER is one of the early international knowledge banks that compiles data, reports and policies for a set of indicators that were developed by the World Bank (World Bank, 2013).

A few comments on the genealogy, proliferation and criticism of international knowledge banks, nowadays sometimes referred to as 'global public goods', are in order here given their eminent role in today's

data-based or evidence-based policy planning (e.g., Global Monitoring Report, 2018). In the education sector, the World Bank has taken the lead in developing and drawing from its knowledge bank to influence national reforms. The concept of an international knowledge bank was first discussed at the Board of Governors of the World Bank in March 1996 (Jones, 2004; Jones with Coleman, 2005). One of the options discussed was whether the financial lending operations should be delegated to the regional development banks (Asian Development Bank, African Development Bank, etc.) while the Bank itself focused on the lending of ideas. Three years later, in 1999, the World Bank's Global Development Network (GDN) was launched at a conference in Bonn (Stone, 2000). The idea was to treat local best practices as a 'public good' and make them globally available. As a result, policy transfer would ideally occur within and among the countries of the Global South, replacing the practice of transplanting reform packages from the First to the Third World.

Although the World Bank has not diminished its role as a money bank, it has acted, over the past decade, increasingly as a global monitor, a lender of 'best practices' and as a research bank or a 'global one-shop-stop for educational knowledge' (Zapp, 2017). Zapp's empirical study of the World Bank's volume and output in terms of policy research is worth mentioning here (2017). A multitude of international organisations – Education for All Fast Track Initiative or the Global Partnership for Education (specifically, the EFA FTI Indicative Framework and the GPE Results Framework), OECD (Education GMS), or UNESCO Paris (Global Education Monitoring Report), to name only a few, have followed suit. These international organisations compare annually national development against an international set of indicators and supplement their data with their portfolio of good practices. In other words, the World Bank has not been alone in constructing and using international knowledge banks to gain leverage at the national level. Used as a governance tool ('governance by numbers'), the ranking and scoring of nations along specific indicators generates far greater reform pressure on ministries of education from low- and middle-income countries and from fragile and conflict affected states than the conventional strategy of imposing loan and grant conditionalities on ministries of finance. In fact, in the wake of these more subtle strategies of inducing reform pressure from within, externally imposed conditionalities appears as a crude and outdated means of inducing change.

The international knowledge banks have been forcefully criticised for a host of reasons (see Steiner-Khamsi, 2007): the covert political agendas

embedded in 'governance by numbers' (Grek, 2008) or in 'epistemic regulation' (Zapp, 2017, p. 2; see also Kamat, 2012), the prescriptive nature of benchmarks and global norm-setting (Steiner-Khamsi, 2016), or the surplus of data (Lubienski, 2019, p. 70) which has created an opportunity for non-state actors (e.g., private sector think tanks) and international organisations (e.g., OECD) to serve as policy brokers that fill the vacuum between knowledge production and policymaking. Similar to digital platforms (LinkedIn, Facebook, etc.), the data and information overload of international knowledge banks has accelerated innovation on one hand (Gawer, 2009) and generated a financially lucrative 'rentiership' business model on the other. As part of this model, global public goods are offered for free, but the interpretation and visualisation of data, consultancies and networking are sold at an economy of scale (Birch, 2020).

That said, a sweeping attack against the very act of collecting comparative data on system variables (class size, teacher remuneration, etc.), system features (tracking of a system, provision of support structures for students with special needs, etc.) and system performance, as done en masse by international knowledge banks, appears to be misplaced. International comparison has become an important policy tool and is not only used by international knowledge banks but also by a wide range of stakeholders in education. Civil society organisations and interest groups, for example, draw on international comparison, rankings and scores to advocate for their cause and to generate reform pressure on decision makers.

Circling back to the meaning of 'system' in sociological system theory, national education systems are considered to be merely 'national organisations' or national segments of the function system of education. For sociological system theory, there exists only one worldwide education system with its own logic, code and mode of regulation. It is important to bear in mind that systems theory does not attempt to be an action theory. Quite to the contrary, Luhmann finds regulatory science, steering or 'the theory of planning' to be 'in a desolate state' (Luhmann, 1997, p. 41) and diametrically opposed to his notion of 'self-steering' of the system. Given that only one function system of education exists, researchers in policy studies often draw their attention to processes of secondary differentiation or segmentation to use national education systems or organisations as their units of analysis. As pointed out astutely by Armin Nassehi (2005, p. 179), any system-theoretical analysis of organisations must begin with why and how questions, such as why organisations tend to describe themselves as rational and efficient. Dissecting self-descriptions

of organisations provides important clues on how a system operates and how it justifies its decisions. In a similar vein, understanding the process of secondary differentiation is essential for policy borrowing researchers as we attempt to understand, among other things, why a global education policy resonates at a national level and how one and the same global education policy is enacted or translated differently at the national level.

KEY CONCEPTS OF SOCIOLOGICAL SYSTEMS THEORY

It is important to bear in mind that Niklas Luhmann's interpretive framework is quintessentially a sociological theory. Luhmann makes in his published work an unabashed commitment to understanding how social systems – rather than how individuals inhabiting these systems – communicate or operate. His is the sociological project of understanding social systems, rather how the 'five billion psychological systems' (individuals) communicate, observe or act (Luhmann, 1990, p. 78, cited in King & Thornhill, 2003, p. 6). Confounding the two units of analysis has apparently been

> an issue that time and again causes difficulty. One can say it a hundred times, but it is always in vain. The observer is not automatically a psychic system; he is not automatically a consciousness. (Luhmann, lecture 7, winter semester 1991–1992, printed in Baecker, 2013, p. 105)

For systems theory, it is not the collective of individuals that makes up a society, but rather the various subsystems with their respective environments: the subsystems of education, politics, law, religion, and so forth. These subsystems observe and differentiate themselves from other subsystems to generate information. All the subsystems of society communicate in their idiosyncratic ways with their own (subsystem-specific) environment and thereby generate shared, societal meaning. Societies, as well as organisations and interactions, operate in an idiosyncratic manner and are, to use systems-theoretical vocabulary, self-referential or autopoietic systems (Luhmann, 1995). The importance placed on how systems generate and reproduce boundaries with their environment makes sociological systems theory quintessentially an ecological theory with a focus on system communication or interaction.

To understand systems theory, readers must also kick the habit of intuitively inserting the adjective 'national' whenever they come across the

term 'the educational system'. As mentioned above, for Luhmann, there is only one universal system of (world) society and only one universal economy, science, legal, political, or education subsystem or function system. This is not to downplay the differences (sometimes vast) in how these different function systems are organised at the national level. In the function system of education, for example, the normative beliefs of how schools should be organised and what should be taught in them vary considerably. Contextual or national differentiation at the organisational level is acknowledged but ultimately of only slight interest to sociological systems theory. Luhmann is primarily concerned with understanding a system's performativity; that is, how it observes and differentiates itself from its environment or other subsystems and how it communicates that difference (Luhmann, 2006).

By definition, a system is at the same time operatively closed and cognitively open. As a result of functional differentiation in modern societies, only the subsystem of education educates, only the subsystem of politics governs, or only the legal subsystem issues laws.

Changes occur as a result of differentiation. In fact, one of the prominent features of modern society is functional differentiation. Society then consists of functionally differentiated systems, all operating with their own codes, identity and modes of regulation. The function systems are closed vis-à-vis other function systems, but they are interdependent. For example, even though the education system has its own education function, its own organisations (schools), its own actors (teachers, students) and its own modes of regulation (until recently driven more by moral and legal rather than market considerations), the function system of education both depends on and contributes to the functioning of other function systems. For example, it depends on the financial resources made available in the economic system.

Three examples should suffice to illustrate the interpretive framework of systems theory. First, a historical account of modern schooling helps to illuminate the process of functional differentiation. A result of the Enlightenment and the nation-state, modern schooling introduced the new social category of a 'school class' (Luhmann, 2002, pp. 119–121, cited in Mangez & Vanden Broeck, 2020) which, encompassed, regardless of social class, all individuals that the state considered to be future citizens of the nation. The discovery of the child as a human species that, regardless of social background, deserves special care and attention helped suspend the earlier focus on the education of princes, noblemen, knights and future sovereigns, taught by private tutors, or first-born sons,

taught by priests. At the same time, modern schooling displaced earlier educative sites, notably the family, the community and the church and, in fact, rendered the other educative sites – including their histories, values, and languages – as illegitimate. In a continuous process of professionalisation and specialisation of education, schools came to be seen as the only organisational form where literacy, numeracy and other relevant skills for living in a modern society were supposed to be taught, thereby disempowering all other educative sites, which were downgraded and framed as non-formal and unprofessional education. Second, systems differentiated themselves by pursuing specific functions and delegating other, unrelated functions to other systems. The function system of education, for example, is to educate students (meant in terms of a social category rather than in an age-specific sense). As Luhmann asserts, the education system neither needs to, nor can generate money, nor exert power, nor produce research for society at large (2002, p. 14). The functions are divided among systems, and all are considered equally important. Third, in a quest to reduce uncertainty, interruption and pertur-bation, a system receives and translates demands for change in a selective or self-referential manner. In other words, function systems observe and react to each other, but are not able to communicate with each other because each one of them is bound by its own code and language of communication. Finally, systems are not only self-referential but also self-reflective and self-aware. Systems observe their environment and, in an effort towards 'boundary maintenance' (Luhmann, 2006, p. 38), differentiate themselves from other systems.

'TRAVELING REFORMS', THEORISED FROM A SYSTEM-THEORETICAL PERSPECTIVE

A book chapter is too short to elaborate on a complex theory. I there-fore need to confine my explanations to a few concepts of sociologi-cal systems theory that are helpful for understanding the fascinating phenomenon of 'travelling reforms'. Such reforms or global education policies surface, with a time lag, in different national education systems and nowadays often propelled by various international organisations or development agencies; sometimes with financial support in the form of grants or loans. Unsurprisingly, the field of comparative and international education has been enamoured with the study of policy transfer, diffusion and borrowing because our investigations often rest on a comparative, cross-national method of enquiry. Furthermore, the expectation that one

may learn 'something' from experiences in other educational systems or 'borrow' innovative ideas from them has been put forward in the past two centuries as the raison d'être for why rigorous comparative studies of educational systems are needed (Cardoso & Steiner-Khamsi, 2017).
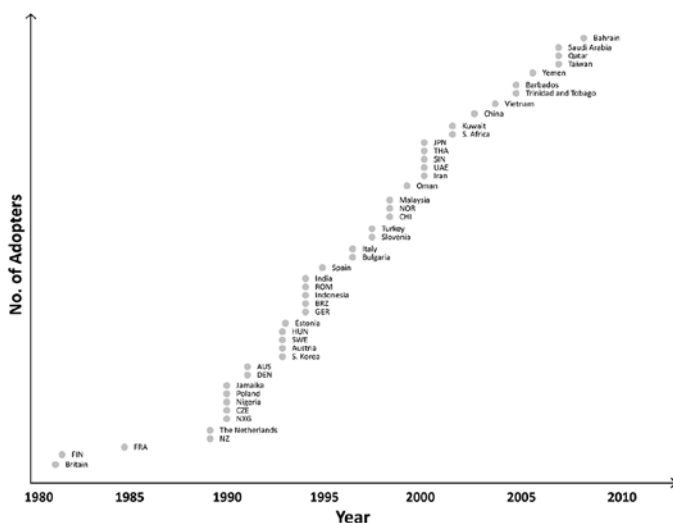
Exploring the Phenomenon that Deserves Theorising

Once we acknowledge that systems are operatively closed but cognitively open, the timing of that cognitive opening up becomes an interesting object of study. Strikingly, the temporal dimension of the global spread of ideas is also widely used in diffusion of innovation studies or critical mass theory. The epidemiological model of diffusion theory traces the deterritorialisation process of a reform over time. It distinguishes between early and late adopters of an innovation. In the early stages, only a few educational systems are 'infected' by a particular reform. Adopters make explicit references to lessons learned from other educational systems, especially those that they are specifically seeking to emulate. At a later stage of explosive growth, however, the policy is globalised or deterritorialised, and the traces to the 'original' are eased. As shown below, the diffusion of innovation model takes on the shape of a lazy s-curve, because at one point the spread of an innovation is saturated.

Several outstanding dissertations in comparative and international education, produced at Teachers College, Columbia University, have applied a system-theoretical lens. Typically, they examine the local policy context to understand why a global education policy has been adopted at a particular time and how it was subsequently translated to fit the local context. Three examples may suffice here: Michelle Morais da Sa e Silva (2017) examined the global spread of conditional cash transfer programmes, Simon Janashia (2015) the spread of per capita financing in the post-Soviet education space, and Rattana Lao (2015) the global dissemination of quality assurance (QA) policies in higher education.

For example, Lao (2015) has produced a fascinating international comparative study on the global diffusion of QA in which she examines in which year higher education systems established formal QA institutions that were separate from ministries of education. As shown in Figure 5.1, her analysis of the higher education literature demonstrates that at least 48 countries had adopted QA policies over the period 1983–2010 QA reforms in higher education. The pioneers were the governments of Britain, France, England, New Zealand and the Netherlands. Starting in the early 1908s, they institutionalised QA by developing distinct policies,

putting mechanisms in place and appointing agencies in charge of QA in higher education. Within the former socialist world system, Poland and the Czech Republic are considered early adopters of quality assessment in higher education. Lao's (2015) analysis resembles the lazy s-curve, depicted by Watts (2003). In line with diffusion of innovation studies, she differentiates between three stages of global reforms: slow growth, exponential growth, burn-out. She identifies the decade of the 1990s as the exponential growth phase of QA. In the new millennium, the adoption of QA is still occurring but at a slower pace; mostly because the higher education landscape is already saturated with QA reforms.

Source: Lao (2015).

Figure 5.1   *The global spread of quality assurance policies in higher education*

The lazy s-curve demonstrates unmistakably why cross-national attraction or bilateral policy transfer represent only one particular case of policy borrowing. For some global education policies, it is more appropriate to assume the existence a global diffusion process or a delta convergence towards the same set of policies across national education systems (Bieber & Martens, 2011). A travelling policy, such as conditional cash transfer programmes (Morais da Sa e Silva, 2017), per capita

financing (Janashia, 2015) or quality assurance in higher education (Lao, 2015) becomes deterritorialised and decontextualised at a take-off point, when several countries adopt the policy or the policy tool, respectively. Eventually, it becomes everyone's and no one's reform at the stage of explosive growth, thereby further increasing its attractiveness to the late adopters.

Key Features of Systems

In the following, I apply four key system-theoretical concepts to the empirical investigation of travelling reforms.

Externalisation

As discussed above, concurrent with operative closure, a system observes other subsystems and communicates about its environment using its own code and logic. In a constant movement between inclusion and exclusion, a system solidifies its identity by means of boundary setting; that is, it distances and thereby differentiates itself from other systems. Interestingly, the system's observation of the difference between itself and other systems also includes the feature of 'second order observation', in that systems observe what other systems observe and therefore observe how the other systems differentiate vis-à-vis their respective environment. Theoretically, the observation of observation may continue ad infinitum, leading to the paradoxical situation that a system observes another's system's observation of the observation of other systems. As will be illustrated later, externalisation generates 'additional meaning' (German: Zusatzsinn) and thereby disrupts the endless circularity of observation and self-observation, or, to use Luhmann's terminology, has a salutary effect on the 'pathology of the paradox' (Luhmann, 1994).

Our preoccupation with the national-level attempts to bring to light the performative act of systems. At particular moments, systems generate national boundaries and reassert themselves as national entities in order to make it appear that there is (global) external pressure for reform or change. In the same vein, they construct other national systems as reference societies at particular moments to suggest that lessons should be drawn from these systems. The political act of externalisation serves to unify or – to use a notion used in policy studies – to build coalitions in support of an educational policy. It is important to point out here that every political act of externalisation necessitates, but also contributes to, the social construction of the nation as an acting subject. Our attention

is directed to how governments deal with policy contestation, at what moments they resort to the semantics of the 'national' and the 'global', and what impact their acts of externalisation have on authorising controversial educational policies in their country.

The group of researchers in comparative and international education that applies the externalisation thesis to the study of school reforms has grown visibly over the past few years (see Waldow & Steiner-Khamsi, 2019). Methodologically, the conceptual focus on externalisation implies an empirical enquiry into how, when and by whom (explicit) references are made. In comparative policy studies, in particular, the investigation entails analysis of which countries are referenced, when a reference is made and how the reference is interpreted. For example, whose publications (including the publications of which organisations) are cited as authoritative texts, which education systems are considered as sources for emulation, within which policy context is a reference to ILSA results (PISA, TIMSS, etc.) made, and how policy analyists in a given country make sense of ILSA results. Several system theorists use the rapid growth of standardised international comparison as an opportunity to understand why PISA resonates and how PISA results are translated or interpreted in various policy contexts. We do not assume that PISA has a priory a salutary effect on national school reform, but we rather analyse why, how and when national policy actors 'open up', or are receptive to the system's 'irritations' caused by PISA, TIMSS (Trends in International Mathematics and Science Study) or by any other ILSA, and how they subsequently translate these external impulses into the language of the organisation or the national policy context.

Projections
The focus on the idiosyncrasies of a system and its national forms of organisation brings a fascinating phenomenon to light that at first sight appears to be contradictory: despite the widespread rhetoric of learning from 'best performing' school systems, there is no universal consensus on why some school systems do better than others on tests such as PISA. On the contrary, there is great variation in how national governments, media and research institutions explain Finland's, Shanghai's or Singapore's 'success' in PISA or TIMSS. However, there is a pattern in these varied, sometimes diametrically opposed, explanations, which is best captured by the term 'projection' (Waldow, 2019; Waldow & Steiner-Khamsi, 2019). The reception and translation of ILSA results reflect controversial

policy issues in a country's own policy context, rather than the actual organisational features of the league-leaders.

'Finnish success' is a good case in point. There is a long list of explanations for why Finnish students do well on ILSAs. Depending on what the controversial policy issue is for which policy actors seek an (internally induced) quasi-external source of authority, Finland's success is alternately attributed to its strong university-based teacher education system, the system of comprehensive schooling with minimal tracking of students, or the nurturing environment in schools where students ironically are exposed to very few high-stakes standardised tests.

The same applies for the league-leaders themselves: depending on the timing, notably whether the positive results are released at the end or the beginning of a reform cycle, the policy actors tend to take credit for the positive results or, on the contrary, belittle the success with the explanation that the students performed well for all the wrong reasons, including private tutoring, stressful school environment, and learning to the test (see Waldow & Steiner-Khamsi, 2019).

Reference and counter-reference societies

In comparative education, the term 'reference society', is key for an epistemological understanding of policy borrowing. At centre stage is the question of which society or which educational system is referenced and used as a model for emulation or lesson-drawing. This body of scholarship is closely associated with studies on 'reference society' presented by sociologist Reinhard Bendix (1978, p. 292). Bendix used the term to denote how governments used economic competitors and military rivals as reference societies for their own development. One of the examples, discussed by Bendix, is the fascination of Meiji-era Japan with the West. In comparative education, the term was first introduced by R. Freeman Butts (1973), associate dean and professor of Teachers College, Columbia University (see Waldow, 2019). Butts observed that governments of developing countries frequently use a specific educational system in the Global North as a model for emulation. That country's path to 'modernisation' served government officials in the Global South as a reference for educational reforms in their country. It is important to bear in mind here that, at the time of Butts, transnational networks and dependencies, established during colonial times, endured into the present and determined in great part the choice of reference societies. In a similar vein, the historian and comparativist David Phillips coined the term 'cross-national attraction' to denote the keen interest of

nineteenth-century British government officials in educational reforms of Germany (Phillips, 2004). Both Butts and Phillips used records of study visits and government reports as sources for their findings on cross-national attraction or policy borrowing. The link between reference society and political change has also been well documented in comparative education research. Examples include two cases of a radical change in reference societies as a result of fundamental political changes in post-Soviet Latvia and post-socialist Mongolia, respectively (see Silova, 2006; Steiner-Khamsi & Stolpe, 2006).

Strikingly, studies on reference societies and cross-national policy attraction have experienced a revitalisation of a special sort in recent years with the fast advance of international large-scale student assessments (ILSAs), used in many countries as a policy tool for governance by numbers. The preoccupation with what league leaders (Finland, Shanghai, Singapore, etc.) have 'done right' has generated a new momentum for policy borrowing research. Precisely at a stage in policy borrowing research when scholars have put the study of cross-national policy attraction to rest and instead directed their attention to the ubiquitous diffusion processes of global education policies in the form of 'best practices' or 'international standards' vaguely defined, the cross-national dimension – and by implication the focus on the nation-state and its national policy actors – has regained importance in ILSA policy research. In the case of PISA, the preoccupation of national policy actors is, at least rhetorically, on how their own system scores as compared to others, and what there is to 'learn' from the league winners, league slippers and league losers, in terms of PISA's twenty-first-century skills. Because policy actors often attribute 'best practices' to particular national educational systems, the national level regained importance as a unit of analysis. ILSA policy researchers therefore found themselves in a position of having to bring back the focus on national systems; a unit of analysis criticised as 'methodological nationalism' which, if used naively, is cause for concern because of its homogenising effects (Wimmer & Glick Schiller, 2003).

Research on reference societies has also been refined over the past few years in other ways. For example, intrigued by negative media accounts in Germany about the PISA league leader Shanghai (during the 2012 PISA round), Florian Waldow scrutinised the policy usage of 'negative reference societies' (Waldow, 2019) or 'counter-reference societies', respectively. The concept of the reference or counter-reference society is based on commensurability. How do national policy actors make the educational systems of league winners appear to be comparable to their own

educational system, in order to suggest that lessons could be drawn? Vice versa, how do they manage to make two educational systems incommensurable and incomparable in order to avoid lesson-drawing? The disbelief or the downplaying, respectively, of Chinese success in ILSAs, notably in the PISA rounds 2012 and 2018, is comparable to earlier, stereotypical accounts of Japanese or pan-Asian education. Similar to the US media accounts of 'A Nation at Risk' (National Commission on Excellence in Education, 1983) in which American policy analysists attempted but ultimately failed to persuade Americans of the great benefits of the German and Japanese educational systems, the education systems of Beijing, Shanghai, Jiangsu and Zhejian are, despite 'PISA success', hardly used as models for emulation in Western countries. As with the 'A Nation at Risk' report, the common reaction to Chinese success reflects a 'yes, but …' attitude (see Cummings, 1989, p. 296): even though there is a general agreement about the outstanding student performance in ILSAs in Hong Kong, Japan, Korea, Macao, Singapore and select cities of PR China, there are too many negative stereotypes associated with education in these locations to assign them reference or emulation status. In fact, the exaggerated statements or myths about 'Asian education' include images of overly ambitious mothers ('tiger mothers'), excessive use of cram schools, competition and suicide among students, elitist higher education, and social inequality. More often than not, the educational systems in Asia are politically instrumentalised as a counter-reference, that is, examples of how educational systems should not develop.

Reception and translation

My preoccupation with externalisation – later captured with the dual term reception and translation – began with the intellectual desire to understand a phenomenon that at first seemed to be odd. In my early work on multicultural education policies (Steiner-Khamsi, 1992), I noticed that some education policies travel from one country to another thereby generating global reform movements. Inspired by the Anti-Apartheid Movement and embraced by the New Left in the United Kingdom, a progressive movement spread at the time within Europe which demanded to drop the lopsided notion of multicultural education (which culturally exoticised disenfranchised minorities and in effect meant assimilation and compensatory education programmes for immigrants) and replace them with more politicised and combative anti-racist policies which targeted the elimination of structural discrimination. I found it fascinating to see that not all school systems were open to this discursive shift and,

even if they were, they implemented the elements of the transnational anti-racist education policy selectively. It did not come as a surprise that policies get implemented differently than planned. What was notable, however, was policymakers' explicit references to other countries. Progressive policymakers insisted that they borrowed the concept even if their own contextualised variant barely resembled the original model that they supposedly had emulated. They were eager to signal their sense of belonging to a larger (Western) European space that tackled discrimination and inequality in schools. What was more: externalisation, or the recourse to Europe or the world, occurred whenever policymakers were under political pressure to justify the introduction of controversial reforms.

As I started to differentiate between diffusion (passive) and reception/translation (active), it became apparent that reforms do not simply travel from one country to another. A more agency-oriented explanation was needed; one that acknowledged the active role of policy actors in importing, borrowing or adopting a travelling reform.

For Luhmann, autopoietic or self-referential systems are meaning-producing entities that make sense out of the (nonsensical) irritations, disturbances and noise which they experience around them:

> In the context of autopoietic reproduction, the environment functions as irritation, as disturbance, as noise, and it only becomes meaningful for the system, when it can be related to the networks of decisions of the system. (Luhmann, 1992, p.173, cited in Vanderstraeten, 2002, p. 246f.)

In the same vein, self-referential systems produce their own causes and determine what counts as internal and external causation (see Fuchs, 1988). External causes – such as the neoliberal move to outcomes-based governance and knowledge-based regulation in public administration – are effective only if they resonate internally, are rendered meaningful and translated into the logic and language code of the system. In education, the orientation towards outcomes was translated into the shift from teaching to learning and from government to governance (Fenwick, Mangez & Ozga, 2014; Mangez & Vanden Broeck, 2020), a move that ultimately resulted in a 'datafication' of learning (Thoutenhoofd, 2018). Whether the class size was large or small, teachers qualified or unqualified, sufficient resources allocated to schools or not (to list only a few input indicators), was rendered irrelevant compared to the question of whether students had achieved the prescribed learning outcomes. The continuous

testing of students became the primary policy tool for monitoring quality and equity in education; that is, rephrased in the terminology of systems theory, a technology for reflection in the system of education.

From the perspective of sociological systems theory, two substantial developments have influenced the type of studies that are currently produced in policy borrowing research: (1) the move from normative to analytical policy borrowing research and (2) the focus on the relation between the local and the global. The first development brings to light the political dimension of cross-national and international references, whereas the second discursive shift draws attention to when, how and why politicians and their advisors make references to 'best practices' or broadly defined international standards to generate a quasi-external reform pressure that helps create political coalitions for their own reform initiatives.

First, the critical perspective of policy borrowing research comes to light in first move, notably, the research interest in how governments justify the need for reform or the need to alleviate reform pressure, respectively. This reorientation towards legitimacy issues has helped comparativists to put the normative and prescriptive orientation of earlier policy borrowing research to rest. In fact, the applied, normative and prescriptive focus on policy learning (what should be learned or borrowed from others?), has been severely attacked in comparative education (e.g., Schriewer, 1990; Steiner-Khamsi, 2004) and replaced with an analytical focus that attempts to understand the receptiveness of a system to experiences from elsewhere.

Second, the relational nature between the global and local is not to be underestimated. One of the most dazzling phenomena is that local politicians periodically invoke globalisation semantically and present the condition of globalisation towards their local audience as a quasi-external force for the sole purpose of generating reform pressure on local policy actors. The fact that a series of similar global education policies, such as, for example test-based accountability (Verger, Parcerisa & Fontdevila, 2019), circumvent the globe is often taken as a proof that national educational systems are converging towards the same reform package or towards the same set of global education policies. Note the circularity of their argument: local politicians first create the phantom of (vaguely defined) international standards to generate reform pressure; then they use the existence of such (self-produced) standards as proof that all educational systems, including their own, must be aligned with them. To put it differently, 'globalisation' is a reality but also a phantom that is period-

ically mobilised for political and economic purposes. The differentiation between globalisation as a condition (real globalisation) and a discourse (imagined globalisation) comes to mind here (see Steiner-Khamsi, 2004). To reframe this statement in terms of sociological systems theory: each function system observes other systems and works to differentiate itself. Because of this observation, each system can externalise to other systems in times of uncertainty when self-referentiality comes across as paradoxical and 'additional meaning' (Zusatzsinn) is needed for its operation. 'Globalisation', or rather the reference to globalisation, is thus an internally induced, meaning-making policy tool that may be instrumentalised to justify policy decisions at national level.

CONCLUSIONS

To reiterate the points made in the previous sections, a multitude of research questions open up when a coherent and complex social theory is applied to explaining a phenomenon; here the phenomenon of travelling reforms: at what particular moments do systems externalise, that is, open up, observe and reference other countries or other (reform) experiences? Whom or rather what do they choose as their object of observation (reference and counter-reference societies)? What do they do with the observation or reference, that is, how do they (back-) translate it to fit their own (system) logic? To sum up, system-theoretical studies typically investigate when systems open up, analyse the idiosyncratic projections made into the other countries' education system, examine which other school systems are selected as reference societies or counter-reference societies, respectively, and finally attempt to understand why exactly a particular global education policy resonates in a particular context or is well received and how it is subsequently recontextualised or translated into the 'socio-logic' of a country's education system. Meant to pique the reader's curiosity, these questions merely represent a small sample of questions that arise when the concepts of system closure, self-referentiality and observation/reference are used as sources of inspiration to explore relevant topics in comparative policy studies.

REFERENCES

Baecker, D., ed. (2013). *Introduction to Systems Theory*. Cambridge: Polity Press (2nd edition).

Bendix, R. (1978). *Kings or People: Power and the Mandate to Rule*. Berkeley: University of California Press.

Bieber, T. & Martens, K. (2011). The OECD PISA Study as a Soft Power in Education? Lessons from Switzerland and the US. *European Journal of Education*, *46 (1)*, 101–116.

Birch, K. (2020). Technoscience Rent: Toward a Theory of Rentiership for Technoscientific Capitalism. *Science, Technology, and Human Values*, *45 (1)*, 3–33.

Butts, R. F. (1973). New Futures for Comparative Education. *Comparative Education Review*, *17 (3)*, 289–294. doi: 10.2307/1186966.

Cardoso, M. & Steiner-Khamsi, G. (2017). The Making of Comparability: Education Indicator Research from Jullien de Paris to the 2030 Sustainable Development Goals. *Compare*, *47 (3)*, 388–405.

Cummings, W. K. (1989). The American Perception of Japanese Education. *Comparative Education*, *25 (3)*, 293–302.

Fenwick, T., Mangez, E. & Ozga, J., eds (2014). *World Yearbook of Education. Governing Knowledge: Comparison, Knowledge-Based Technologies and Expertise in the Regulation of Education*. London and New York: Routledge.

Fuchs, S. (1988). Translation and Introduction of 'Tautology and Paradox'. In N. Luhmann (ed.), *The Self-Descriptions of Modern Society* (6th edition, pp. 21–37). Sociological Theory.

Gawer, A. (2009). *Platforms, Markets and Innovation*. Cheltenham: Edward Elgar Publishing.

Global Education Monitoring Report. (2018). Fulfilling our Collective Responsibility: Financing Global Public Goods in Education. *Policy paper 34*. Paris: UNESCO Global Education Monitoring Report.

Grek, S. (2008). From Symbols to Numbers: The Shifting Technologies of Education Governance in Europe. *European Educational Research Journal, 7 (2)*, 208–218.

Janashia, S. (2015). *Introduction of Per-Capita Education Financing in Former USSR Countries*. New York: Teachers College, Columbia University, Proquest Digital Dissertations.

Jones, P. W. (2004). Taking the Credit: Financing and Policy Linkages in the Education Portfolio of the World Bank. In G. Steiner-Khamsi (ed.), *The Global Politics of Educational Borrowing and Lending*. New York: Teachers College Press, pp. 188–200.

Jones, P. W. with D. Coleman (2005). *The United Nations and Education. Multilateralism, Development and Globalization*. London and New York: Routledge.

Kamat, S. (2012). The Poverty of Theory. The World Bank's System Approach. In Steven J. Klees, Joel Samoff & Nelly P. Stromquist (eds), *The World Bank and Education. Critiques and Alternatives*. Rotterdam: Sense, pp. 3–20.

King, M. & Thornhill, C. (2003). *Niklas Luhmann's Theory of Politics and Law*. New York: Palgrave.

Lao, R. (2015). *A Critical Study of Thailand's Higher Education Reforms: The Culture of Borrowing*. London: Routledge.

Lubienski, C. (2019). Advocacy Networks and Market Models for Education. In M. Parreira do Amaral, M. Steiner-Khamsi & C. Thompson (eds), *Researching the Global Education Industry*. Palgrave, pp. 69–86.

Luhmann, N. (1990). *Essays on self-reference.* New York: Columbia University Press.

Luhmann, N. (1992). Organisation. In W. Küpper & G.Ortmann (eds), *Mikropolitik: Rationalität, Macht und Spiele in Organisationen*. Opladen: Westdeutscher Verlag (2nd edition), pp. 165–185.

Luhmann, N. (1994). *Vorsicht vor zu raschem Verstehen. Talk-Show mit Niklas Luhmann über Entscheidungsvermögen*. Interviewer: Alexander Kluge. dctv. tv. Retrieved from: https://www.dctp.tv/filme/news-stories-04-07-1994

Luhmann, N. (1995). *Social systems.* Stanford, CA: Stanford University Press.

Luhmann, N. (1997). Globalisation or World Society: How to Conceive of Modern Society? *International Review of Sociology, Revue Internationale de Sociologie, 7 (1)*, 67–79.

Luhmann, N. (2002). *Das Erziehungssystem der Gesellschaft*. Frankfurt/M: Suhrkamp.

Luhmann, N. (2006). System as Difference. *Organization, 13 (1)*, 37–57.

Mangez, E. & Vanden Broeck, P. (2020). The History of the Future and the Shifting Forms of Education. *Educational Philosophy and Theory*, *52 (6)*, 676–687. https://doi.org/10.1080/00131857.2019.1708329

Morais da Sa e Silva, M. (2017). *Poverty Reduction, Education, and the Global Diffusion of Conditional Cash Transfers.* New York: Palgrave.

Nassehi, A. (2005). Organizations as Decision Machines: Niklas Luhmann's Theory of Organized Social Systems. *The Sociological Review, 53 (1)*, 178–191.

National Commission on Excellence in Education. (1983). A Nation at Risk: The Imperative for Educational Reform: A Report to the Nation and the Secretary of Education, United States Department of Education. Washington, D.C.: National Commission on Excellence in Education.

Phillips, D. (2004). Toward a Theory of Policy Attraction in Education. In G. Steiner-Khamsi (ed.), *The Global Politics of Educational Borrowing and Lending*. New York: Teachers College Press, pp. 54–67.

Schriewer, J. (1990). The Method of Comparison and the Need for Externalisation: Methodological Criteria and Sociological Concepts. In J. Schriewer in cooperation with B. Holmes (eds), *Theories and Methods in Comparative Education*. Bern: Lang, pp. 3–52.

Silova, I. (2006). *From Sites of Occupation to Symbols of Multiculturalism. Reconceptualizing Minority Education in Post-Soviet Latvia*. Information Age Publisher.

Steiner-Khamsi, G. (1992). *Multikulturelle Bildungspolitik in der Postmoderne* [Multicultural Educational Policy in Postmodernity]. Opladen: Leske & Budrich.

Steiner-Khamsi, G., ed. (2004). *The Global Politics of Educational Borrowing and Lending.* New York: Teachers College Press.

Steiner-Khamsi, G. (2007). International Knowledge Banks and the Production of Educational Crises. *European Educational Research Journal, 6 (3)*, 285–293.

Steiner-Khamsi, G. (2016). Global Indicators and Local Problem Recognition: An Exploration into the Statistical Eradication of Teacher Shortage in the Post-Socialist Region. In Karen Mundy, Andy Green, Bob Lingard & Antoni Verger (eds), *The Handbook of Global Education Policy*. Chichester: John Wiley and Sons, pp. 573–589.

Steiner-Khamsi, G. & Stolpe, I. (2006). *Educational Import in Mongolia: Local Encounters with Global Forces*. New York: Palgrave Macmillan.

Stone, D. (2000). *Banking on Knowledge: The Genesis of the Global Development Network*. London and New York: Routledge.

Thoutenhoofd, E. D. (2018). The Datafication of Learning: Data Technologies as Reflection Issue in the System of Education. *Studies in Philosophy and Education, 37*, 433–449. https://doi.org/10.1007/s11217-017-9584-1

Vanderstraeten, R. (2002). The Autopoiesis of Educational Organizations: The Impact of the Organizational Setting on Educational Interaction. *Systems Research and Behavioral Sciences, 19 (3)*, 243–253. https://doi.org/10.1002/sres.481

Verger, A., Parcerisa L. & Fontdevila, C. (2019). The Growth and Spread of Large-Scale Assessments and Test-Based Accountabilities: A Political Sociology of Global Education Reforms. *Educational Review, 71 (1)*, 5–30. doi: 10.1080/00131911.2019.1522045.

Waldow, F. (2019). Introduction: Projection in Education Policy-Making. In F. Waldow & G. Steiner-Khamsi (eds), *Understanding PISA's Attractiveness: Critical Analyses in Comparative Policy Studies*. London: Bloomsbury, pp. 1–13.

Waldow, F. & Steiner-Khamsi, G., eds. (2019). *Understanding PISA's Attractiveness: Critical Analyses in Comparative Policy Studies*. London: Bloomsbury.

Watts, D. J. (2003). *Six Degrees: The Science of a Connected Age*. W. W. Norton & Company.

Wimmer, A. & Glick Schiller, N. (2003). Methodological Nationalism, the Social Sciences, and the Study of Migration: An Essay in Historical Epistemology. *International Migration Review, 37 (3)*, 576–610. doi: 0198-9183/03/3703.0143.

World Bank. (2013). *SABER. The What, Why, and How of the Systems Approach for Better Education Results* (Conference Edition). Washington, DC: World Bank.

Zapp, M. (2017). The World Bank and Education: Governing (Through) Knowledge. *International Journal of Educational Development, 53*, 1–11.

6. Using a systems approach to education and development: insights from a multi-country research programme on access and learning

Keith M. Lewin

INTRODUCTION

Education systems lie at the heart of development and access to education is both a part of the definition of poverty and a means for its reduction. Since the 1960s development agencies have prioritised investment in education as one of the primary vectors through which to accelerate economic development, promote social mobility out of poverty and democratise governance. Educational services are delivered through school systems that are largely publicly financed and shaped by national and individual aspiration, physical infrastructure, pedagogic preference, curricula choices, high-stakes selection examinations, resource availability, and the roles and responsibilities of administrators, educators and learners. Systems approaches can help frame possibilities and probabilities for development strategies that shape learning opportunities.

This chapter presents a way of thinking about education systems developed by a large-scale Department of International Development (DFID) programme of research initially commissioned in 2005. This built on a long tradition, dating from the 1960s and before, of using systems thinking of one kind or another to understand educational development and identify policy options and opportunities in low-income countries (e.g., Anderson & Bowman, 1965; Halsey, Floud & Anderson, 1965; Coombs, 1968, 1985; Dore, 1976; World Bank, 1980; Colclough

& Lewin, 1990; Lockheed & Verspoor, 1991; Grindle & Thomas,1991; Lewin & Stuart, 2003).

The Consortium for Research on Educational Access, Transitions and Equity (CREATE) worked across four core countries with research teams in national institutions and high-level advisory groups of key stakeholders. CREATE initiated its programme of research by commissioning system-level country analytic reviews undertaken by research teams embedded in the education systems in India, Bangladesh, Ghana and South Africa led by prominent national researchers, with sub-studies in six other countries. These analytic reviews were critical to the research and were launched at national workshops. They were a springboard for studies of different facets of educational access and meaningful learning. CREATE involved over 100 researchers and produced over 200 journal articles, monographs, books, policy briefs and research reports.[1]

This chapter focuses on a small part of CREATE's research and only one aspect of its systems thinking. The basic CREATE model identifies zones of exclusion which condition access to learning and which shape transitions for learners as they progress through education systems. Changing patterns of access to learning illustrate the dynamics of system evolution. The CREATE approach provides a counterpoint to narratives about access, participation, learning outcomes and exclusion based on aggregated data, located at single points in time, and removed from cultural and organisational context. Cross-sectional snapshots can mislead and may generate ineffective interventions detached from system histories and demonstrated capabilities especially where aspects of development are nonlinear. CREATE provided opportunities to reflect on the political economy of transitions and reform on the basis of the evidential record rather than aspirational ambition. It deliberately took the view that global strategies can provide frameworks but are not blueprints and that educational reform is more system specific than it is generic (Little, 2008; Lewin & Little, 2011; Little & Lewin, 2011).

The United Nations Sustainable Development Goal 4 (SDG4) specifies global targets for education systems. These include universal access to preschool through to grade 12 (UNESCO, 2016) coupled to minimum levels of learning achievement. Modelling based on insights from CREATE using data from low- and low-middle-income countries

---

[1]   These are available at www.create-rpc.org and summarised in Lewin (2011a) and in Lewin (2015).

(LICs and LMICs) demonstrates the impossibility of the goals being achieved by 2030. Universal participation would require massive enrolment increases, especially in secondary and pre-school grades, of several times or more students over current levels. No imaginable amount of external assistance would be sufficient to finance the additional recurrent costs of teachers' salaries. Demonstrably most LICs and many LMICs cannot finance SDG4 from domestic revenues without substantial fiscal reform (Lewin, 2017). In addition, current levels of learning are standard deviations below what would signify achievement in line with national curricula and international benchmarks.

SDG4 is much more ambitious than previous agendas (Colclough & Lewin, 1990). These were linked to viable financing. Some advocacy is now privileging hyperbole over reality. These issues are central to current debates on the future of aid to education (Heyneman & Lee, 2016; Burnett, 2019; Lewin, 2020a; Beharry, 2021). The kind of systems analysis CREATE has supported is essential to improve evidence-based allocative decisions that increase the chances of identifying a more balanced and sustainable set of goals for education system development that have some prospect of being achieved.

Following this introduction, section 2 of this chapter describes the CREATE model of access to education. This profiles zones of exclusion and inclusion. It provides a flexible and resilient tool for analysis and policy dialogue. It is presented in two forms – one which will be familiar to planners used to analysing EMIS data, and the other which maps more qualitative aspects of the flow of students through education systems and across transition points. The invitation is to use the model with data from specific education systems. The third section describes an expanded vision of access linked to 'meaningful learning' to promote reflection on the kinds of learning that transform the development of individuals and societies which goes beyond the self-evident importance of foundational literacy and numeracy. A true 'theory of change' for education systems needs to be a theory of development. Five arenas that influence how education systems change and how learning takes place are identified. The fourth section analyses data on the evolution of patterns of participation by grade in a selection of lower-income countries. This shows how systems evolve with different patterns of participation and exclusion in ways invisible to conventional analysis. The fifth section analyses data on the enrolment of boys and girls to illustrate how mapping gender participation by grade adds to understanding of key issues. The final section

collates insights which point the way to future systems analysis that makes use of system dynamics linked to plausible capabilities.

THE CREATE EDUCATION SYSTEM MODEL OF ACCESS AND LEARNING

CREATE developed a systems-level approach that combined an understanding of flows of children and youth through educations systems with the profiling of 'zones of exclusion'. These zones are characterised by lack of access to education, 'silent exclusions' with enrolment but low learning, transitions from being included to being excluded, and from being enrolled to being out of school. The conceptual tools developed by CREATE provide a cluster of mechanisms to explore changing patterns of access, understand bottlenecks that throttle back enhanced participation, and moderate the transitions that may be more or less equitable in relation to household income, location and differential enrolment of boys and girls.

Policy dialogue depends on developing a common language to describe how systems evolve and the dynamics of different types and levels of educational exclusion. This is important for closed systems where the relationship between inputs and outputs is highly predictable. It is even more important for open systems where causal links are less certain and replicable, key parameters change over time, and observations and insight feedback into policy with multi-directional causalities. Education systems have elements of both closed and open systems both of which have to be understood by analysts. Only then can policy be linked to outcomes dynamically, and to the political economies of economic development, social mobility and developmental well-being.

The simplest way of mapping progression through a formal education system is to chart sequential movement of children through grades. This is represented in Figure 6.1. The x-axis plots ten grades of schooling and the y-axis plots the amount of participation at different grade levels.

Zone 0 captures those included and excluded from pre-school. One year of pre-school is now an agreed global goal. Indicators of school readiness of children entering grade 1 are now figuring in national assessments. Some children have no access to pre-school provision. Others have pre-school provision available but are unable to participate for reasons of cost, location, culture or parental preference.

Zone 1 contains those of primary school age who never attend school. It includes those who could attend existing schools but do not, and those
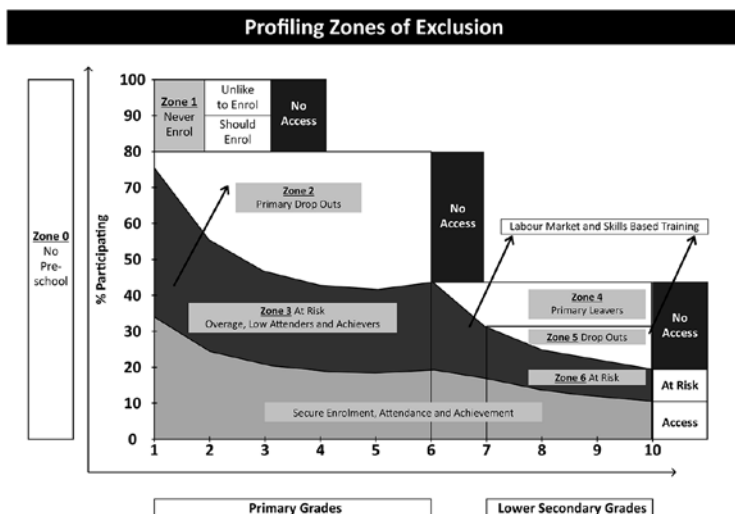
Source: Author (Lewin, 2007a, Lewin & Akeampong, 2009).

Figure 6.1   Profiling zones of inclusion and exclusion

who are excluded by livelihoods, location, civil status, disability, social stigma or other vulnerabilities. Those who have not enrolled by the age of ten years are unlikely ever to enrol.

Zone 2 includes children excluded from education after entering grade 1. These children drop out of primary school. The reasons are well researched and include costs, low achievement, bullying, corporal punishment, violence against boys and girls, and unfriendly schools. As enrolment rates increase the proportion of out-of-school children who are drop-outs, rather than never enrolled, increases.

Zone 3 identifies those in primary school but at risk of drop-out. These children can be described as 'silently excluded' since they are formally enrolled but may learn little, attend irregularly, and be over age (Lewin, 2007a). Those not learning at age-grade appropriate levels become increasingly distanced from the curriculum. Low achievement is widely correlated with drop-out. So is becoming over age in grade, especially where monograde pedagogies make little allowance for variations in capability.

Zone 4 describes those who fail to transit to secondary education as a result of failing to be selected in end-of-primary school selection tests, being unable to afford costs to household which are often three or four times greater at secondary level, problems of travel and security when schools are located at a distance, inequitable gendered exclusions and lack of facilities to address disadvantage and disability.

Zone 5 profiles those dropping out of secondary grades for reasons that go beyond those relevant to primary level. Older children have agency and may reject curricula they see as irrelevant. The opportunity costs of schooling can be substantial in areas where employment is available. Early marriage and pregnancy may also result in premature drop-out.

Zone 6 defines those at risk of drop-out from secondary school and likely to be in one or other of the subcategories of 'silent exclusion'. These include being nominally enrolled but irregularly attending, falling a grade or more behind necessary learning levels, socially or linguistically excluded by cultural group norms or medium of instruction, and isolated by poor management of adolescence.

The model of flows of children expressed on a two-dimensional graph will be familiar to educational planners who work with school census data on enrolments and who use cross-sectional snapshots of enrolment patterns to address policy issues. When the data is viewed over time it adds a dimension to understanding of how patterns of development are unfolding and indicates the direction of travel of system development. At the simplest level children at different nominal grade levels are 'included' if they are registered and are therefore included in school census data. Iterations of data can redraw the chart with variations that suit different sets of concerns. Thus the model can also be populated using separate data for girls and boys, for children from households in different wealth quintiles, by disability status and by affiliation with different cultural groups. The model can be extended up to grade 12 or higher.

Total enrolment by grade is the simplest measure of participation and is easily understood. Grade-by-grade enrolment rates (number enrolled in the grade divided by the number in the appropriate age cohort) can be used to indicate the proportion of children enrolled. However this indicator introduces uncertainties from census data about the size of single-age cohorts. These are often derived from projections over ten-year periods. Attendance rates cannot be used usually because of poor availability in a consistent form. Completion rates for each grade are another option but are even more difficult to measure reliably. The numbers of candidates sitting for terminal grade public examinations can indicate successful

participation but need careful interpretation to account for repeaters and over-age students. In sum total enrolment by grade is the simplest option and is widely available. Grade-specific enrolment rates are better if the data on which they are based are reliable.

Estimating the proportion of children silently excluded at each level requires system specific data and judgement. Flow charts can incorporate data on attendance and time on task, age in grade slippage, and on levels of achievement (Lewin & Akyeampong, 2009; Zeitlyn & Hossain, 2011). Children who are enrolled but learning well below grade expectations are invisible to simple indicators of enrolment. They may constitute a large minority, or even a majority, of learners. The CREATE model can be used to identify what proportion of a particular grade group is both nominally enrolled and achieving at an appropriate level for their grade. This has to be based on national curriculum tests that assess what is taught in the high-stakes assessments that matter to most candidates. This avoids heroic assumptions about curriculum convergence across countries that are unconvincing. It makes it unnecessary to make comparisons across countries that depend on benchmarking against the performance of the highest scoring country as the end in view. Neither is it necessary to assume that performance on learning assessments progresses linearly with grade on the same scale across countries and cultures as is the case if Learning Adjusted Years of Schooling (LAYS) is used (Filmer et al., 2018). Children often do not grow linearly either physically or cognitively (Lewin, 2020b).

Figure 6.2 presents the flow of children through education systems in a different way. This iteration of the model includes pre-school and early childhood care as part of a full picture of educational access and meaningful learning through to grade 12. This qualitative style of presentation highlights categories of exclusion rather than illustrates the magnitude of flows. It can be used for policy dialogue around different transitions and potential points of intervention.

Critically Figure 6.2 highlights how transitions can be internal (e.g., from primary to secondary, transfers from one school location to another) or external (becoming a drop-out, leaving school to enter into a livelihood or into employment (Laugharn, 2007; Hunt, 2008; Ananga, 2011)). Transitions are often linked with discontinuities in enrolment, attendance and achievement. Flows may be regulated more by opportunities at the next level of the education system than by the rights of learners to continue learning or needs to increase the endowment of human capital in the population. Flows ultimately shape mobility linked to educational

Zones of Exclusion and Transitions from Child to Adult

©k.m.lewin@sussex.ac.uk
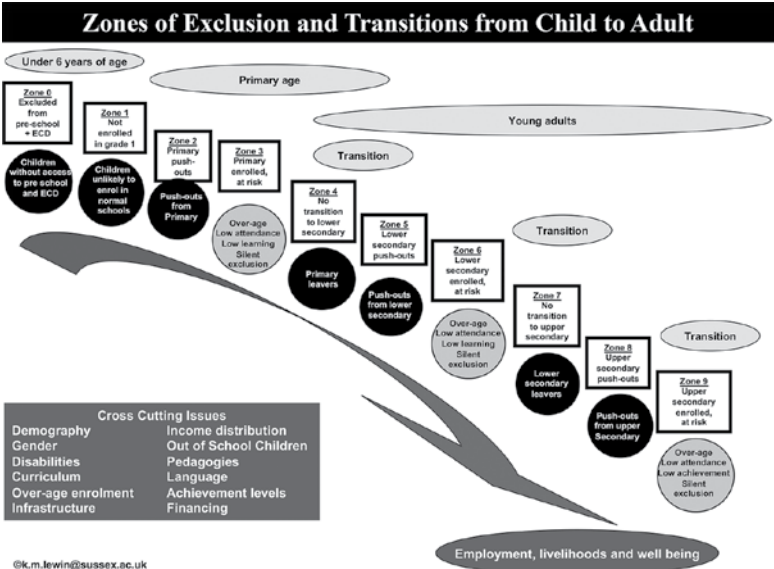
Source: Developed from Lewin & Little (2011).

Figure 6.2   *Zones of inclusion and exclusion from pre-school to grade 12*

levels and determine 'Who gets what' (Lewin & Sabates, 2011) in terms of educational opportunities at higher levels.

Amongst the most critical of transitions are those into livelihoods and jobs. Open systems respond to signals of effective demand and competition for access to jobs rationed by educational qualifications. Monitoring flows through education system has to be juxtaposed with an understanding of the transitions that channel young people into further education and training, employment and livelihoods. The shadow of informal labour markets is present in upper primary schools where many school children are of working age. The 'Fourth Industrial Revolution' promises structural shifts that will rewrite the map of the basis arithmetic of youth unemployment for post-primary completers. The perceptions and realities of labour markets and livelihood opportunities feed back into choices of parents and children at different levels and influence the allocative choices they make. As systems expand real and perceived

rates of return to education are likely to fall as demand in small, modern sectors of employment saturates and knowledge-led economic growth outpaces labour force growth. Increased competition for scarce opportunities linked to qualifications that open the door to modern sector jobs will increase the public and private price of selection without necessarily changing who is selected. This is evident from systems analysis of flows (Lewin, 2008).

These and other cross-cutting issues, e.g., demography, should take the gaze of planners outside the imaginary boundaries of closed education system theory. Learning that has utility is at the epicentre of development. Figure 6.2 implies that education for sustainable development has system-level horizons far beyond foundational literacy and numeracy. Open systems cannot be planned mechanistically because they are in large part open systems. The evolution of changing circumstances and levels of goal achievement across countries suggests the need for equifinality – more than one pathway to defined goals – and multifinality – the valuing of more than one goal. This breaks the impasse generated by attempts to identify singular international priorities that can shape convergent global agendas for educational development. These are of diminishing utility as analysis of open system planning foregrounds endogenously determined priorities and nationally owned and financed sustainable development programmes.

EXPANDING THE VISION AND PROFILING EXCLUSION

A rhetorical critique has developed that global summits (e.g., Jomtien in 1990, Dakar in 2000, Incheon in 2015) and the consequent development assistance related to 'Education for All' valued access to education over learning and that what was needed was 'Access + Learning' (LMTF, 2015). Those who participated in the conferences and those who read the communiques know this is not true. Thus the World Declaration on Education for All in 1990 led to the Framework for Action to Meet Basic Learning Needs (UNESCO, 1990). Article 1 was 'Every person – child, youth and adult – shall be able to benefit from educational opportunities designed to meet their basic learning needs'. At the World Education Forum in Dakar in 2000 and at the Incheon WEF in 2015 Article 1 was reaffirmed and the consistent commitment to learning could not have been clearer. If increases in enrolment have overshadowed the need to enhance learning outcomes this reflects the preferences of the imple-

menting agencies rather more than it reflects the aspirations of the flagship declarations (Collier, 2018). If donor-led 'sector wide approaches' did not value learning where does the responsibility lie (Ahmed, 2011)?

From the outset CREATE developed and applied an Expanded Vision of Access and Learning (EVAL) that made it impossible to disentangle participation from learning outcomes (Lewin, 2011a). This deliberately linked physical access to epistemic and pedagogic engagement and to learning and skills that had utility for development. CREATE developed an expanded vision of access designed to go beyond narrow conceptions of access and participation (Lewin, 2011b) to capture some of the aspirations mandated by UN conferences. These implied that access to education has to include judgements of educational quality and process (what children have access to) and of educational outcomes (what competencies and capabilities are acquired and how they are valued). Quality is at least as important as quantity (Lewin, 1985, Alexander, 2008).

In its simplest form (Box 6.1) the CREATE Expanded Vision of Access and Learning (EVAL) has seven separate elements arising from consideration of desirable components that were potentially realisable, consistent with better practice as judged by experienced professionals, and easy to understand at different levels (Lewin, 2011a). Though some elements of the expanded vision can be quantified others can only be judged in context. This may not suit mechanistic systems analysts accustomed to closed system algorithms with highly predictable causalities. It does sit more easily with open systems analysts who have a professional development orientation that values interventions that are iterated by educators based on insights from reflective practice. The expanded vision provides a framework of expectations which can be modified to suit context and used to build a consensus on access with learning at different levels for different purposes.

BOX 6.1   CREATE EXPANDED VISION OF ACCESS AND LEARNING (EVAL) (2011)

1. Access to pre-school and higher levels of schooling at affordable costs.
2. Local access to safe schools with appropriate levels of staffing, learning materials, services and facilities which provide a positive learning environment.

3. Admission and progression through primary and secondary school within a year of the nominal age-in-grade or multigrade provision.
4. Consistent attendance throughout the school year at least 90% of school time.
5. Learning outcomes that have utility and achievement aligned with national norms.
6. Appropriate access to education and training at secondary and higher education levels.
7. Equitable access to affordable schools and colleges of adequate quality.

We now turn to the factors that shape educational exclusion (and inclusion). A straightforward model identifies five different sources of influence over educational outcomes that leads to meaningful access that has utility (Figure 6.3). Gaining insights into the interplay amongst these factors is central to policy dialogue that balances different perspectives on how systems function, assumptions that are made about pathways of causality, and the plausibility of theories of change.
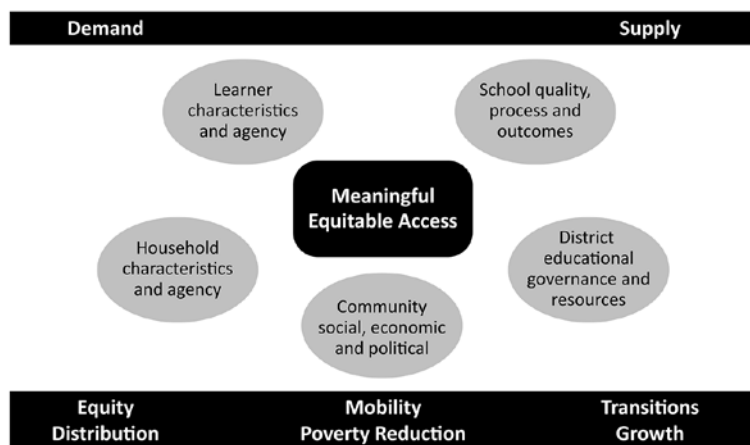
Source: Author's own chart (2007).

Figure 6.3   Model of meaningful and equitable access

This model maps entry points into analytic studies that can give insight into different opportunities for intervention. Each of the arenas identified suggest different units of analysis, i.e., individuals, households, communities, local governance and schools. All these arenas are culturally located and embedded in national systems of governance that extend beyond but interact with education systems. Taking each of the arenas in turn we can note:

1. *Learners* are individuals with attributes that are intrinsic to their identity. They have capabilities that reflect natural endowments, environmental influences, nurture and interactive effects. Beliefs about characteristics of individuals are widespread and are influential in determining many aspects of education policy, e.g., streaming and grouping of children by capability, adopting automatic promotion, identifying 'gifted' children, diagnosing and intervening for low achievers, differentiating curricula for boys and girls, adapting pedagogy to different age groups, and preferences for learning styles. These beliefs are culturally located and embedded in pedagogic traditions.

2. *Households* are at a level above the individual and for many learners are the reference group that influences motivation, provides support for learning, manages opportunities for learning inside and outside the household, and mobilises the resources needed (time, money, learning materials). Household characteristics are strongly associated with participation, progression, completion and learning levels in most education systems.

3. *Community* characteristics are above the level of the household and are an aggregation of common resources, norms of expectation and behaviour, aspirations for development for younger learners being socialised into adult patterns of behaviours and roles in relation to civil society, livelihoods and employment, and health and well-being. Education systems interface with different communities and through collective attendance at events and governance structures that include community representatives.

4. *Governance* of education systems is stratified in ways that resonate with local government structures. It is usually at this level that collective functions are enacted – e.g., enrolment planning, inspection, quality assurance, examinations, teacher employment and deployment across schools, student flow, school location. This is also the

level at which public financing is translated into services to deliver educational outcomes.

5. *Schools* are the symbolic heart of public education systems. They host and curate the learning activities of most young people when education systems are working efficiently and effectively. Schools exist within systems with a constant flow across boundaries of people (teachers, students), things (learning materials, curricular guidelines) and information (enrolment patterns, examination results). In schools learning is enacted and managed. Schools may have more or less autonomy in relation to curricula, pedagogy, and admission and enrolment.

All five arenas identified deserve consideration. If a systems approach is adopted they all interact with each other and with national and international development strategies. The mechanical approaches of closed system theory are of limited application in the real world because the articulation between elements is as much social and political as it is mechanical. Simple principal agent theory assumes 'one actor (the principal) wants a task accomplished, so they employ another actor (the agent) to complete the task' (Spivack, 2021). The CREATE model is a reminder that there is no single principal – actors in all five arenas can lead on defining tasks – and there is no necessary reason why the tasks will be aligned if they arise independently from each arena. Similarly, the agent side of the assumed managerial dichotomy is likely to have many actors all of which somehow must share goals and motivations and reinforce rather than compete with each other for operating responsibilities, resources and infrastructure. Ambiguities permeate this kind of analysis with principals sometimes being agents and agents sometimes being principals and each influencing the other. Schrodinger's cat would have understood.

   Open systems are dynamic and are subject to many exogenous influences that are unpredictable. Insights into one system in one place and at one time are unlikely to translate reliably to others without system specific understandings of probable causalities, interactions, preferences and resource constraints. From an open systems perspective actors in the five arenas have an impact on both the supply and the demand side of learning. Thus on the demand side learners may have preferences for enrolment, progression to higher levels and subject specialisation relatively independent of household characteristics. These co-exist with demand-side choices that are influenced by household level dispositions, local labour

markets and cultural practices. The supply of educational opportunities is widely determined by community and local government-level investment and policy on enrolment and progression (Cameron, 2011). This influences school admissions and the quality of provision which itself may also have an effect on demand. The interactions are complex and contain many feedback loops.

Many externally supported programmes to universalise participation and enhance learning have concentrated on supply-side inputs. This is critical where infrastructure is weak, buildings and classrooms inadequate or unavailable, learning material in short supply, and teacher qualification is low. These inputs are greatly needed where enrolment growth is strong. But opportunities to learn may also be affected by falling demand, especially amongst older children who are the most likely to be out of school and in communities where the opportunity costs of school attendance are high, and where school quality is low. Uneven school quality compromises equitable access to the extent that in some low-enrolment countries the poorest children have less than a tenth the chance of the richest of completing secondary schooling. Open systems approaches capture the importance of both supply- and demand-side aspects of learning and their interactions across system boundaries.

THE EVOLUTION OF PARTICIPATION

The CREATE research explored access and learning across low- and low-middle-income countries using a set of conceptual tools described here and at www.create-rpc.org and in Lewin (2015, 2017). The analysis presented below illustrates the power of just one type of systems analysis based on available data.

Patterns of Participation by Grade in Eight Countries

Administrative data from school censuses allow time series of enrolments by grade to be plotted. Figure 6.4 shows the results for eight countries, indicating how enrolments have changed since 2005 in each grade over the basic education cycle. The numbers enrolled by grade are used for y-axis values to simplify this illustration since these are readily

understood and widely available.[2] Children in a particular grade would normally be expected to be promoted to the next grade the following year. A high-enrolment low-attrition system would therefore show lines on a graph for each year as parallel and close together. The reduction in enrolment in different grades for the same year would decline at the same rate as the change in size of the relevant age cohort. Figure 6.4 demonstrates this is not what has been happening.

First we note that the dotted line running across the charts is a line estimating the number of children in a grade-age group. If this has a negative gradient it indicates that the population of school-age children is growing so each age group is larger than the previous one. This is true in the case in Malawi or Ethiopia. Conversely if the gradient is positive there are fewer children in each age group. This occurs where there is demographic transition as in Vietnam or Cambodia. If the population of children is shrinking it is easier to invest more per child and enhance quality since the number of children relative to the taxpaying working population is falling.

Second, every system has a tipping point where the number of children enrolled is about the same as the number in the age group related to that grade. The tipping point is around grade 3 in Pakistan and grade 8 in Kenya showing there is wide variation between systems. This does not mean that all children are enrolled below the level of the tipping point since enrolments are likely to be inflated by repeaters and overage entrants. Thus Uganda and Malawi have many more enrolled than there are in the relevant age groups below grade 6. If all children entered and progressed at the appropriate nominal age then there would be no tipping point and the enrolment curve would follow the age grade cohort line.

Third, most of the countries in Figure 6.4 have many more children in grade 1 than there are six-year-olds in the population. The excess numbers can be very large and may be double the number of primary school entry-age children in the population. Alarmingly, although enrolments have increased in all the countries selected except Vietnam and Cambodia which have had demographic transition, the over-enrolment in lower grades has generally persisted (e.g., in Malawi, Uganda and Mozambique). This is only evident from systems flow analysis which

---

2   Compound indicators of participation, e.g., gross and net enrolment rates, primary completion rates, etc., introduce more uncertainties of interpretation and measurement than simple enrolments.
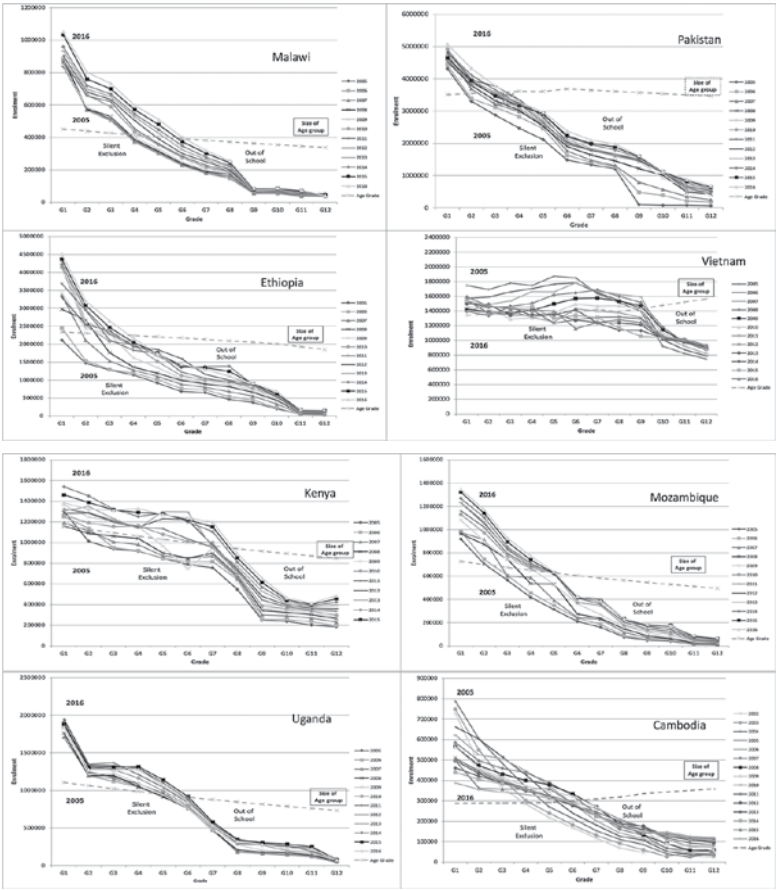
Figure 6.4   Enrolment flow charts from eight countries

draws attention to changing patterns of demand for places at different grade levels. Counter-intuitively, as systems become more efficient and repetition is reduced, the number of places needed can fall as enrolment rates increase.

Fourth, most systems have strong inflexion points in the flow around high-stakes selection points. This can be seen where there is a sudden drop in enrolments between grades. This occurs in Malawi, Ethiopia

Uganda and Kenya, and in Pakistan. Some bottlenecks persist despite overall enrolment increases, e.g., in Malawi the numbers enrolled in grade 9 have been consistently about 30% of those enrolled in grade 8 since 2000, indicating the impact of a high-stakes primary school leaving examination on flows towards grade 12. In Kenya there is a sharp drop from grade 8 to grade 9 as children transit into secondary schools (Oketch & Somerset, 2010). There is a similar drop in Uganda after grade 7. These discontinuities can be exacerbated by queuing whereby children repeat examination grades whilst attempting to improve achievement to be selected to the next level. Bottlenecks that inflate enrolments in examination entry grades are a source of considerable inefficiency.

Figure 6.4 draws attention to the levels at which there are most out of school (OOSC) children. A first order indication of OOSC is provided by the gap between the age grade group line and the enrolment lines. This does not account for all OOSC since over enrolment in lower grades conceals non-attenders and those who have dropped out. In addition, a proportion of enrolment above the tipping point is of over-age children and repeaters. A full accounting needs household data of quality rather than administrative data. However, the charts are simple to construct and do indicate important dynamics. The number out of school is unlikely to be less than the gap between the age group line and enrolments. Children in this space are mostly older not younger.

GEMR data shows that 50% of OOSC are over the age of 15 years and that more than 25% are between 13 and 15 years. Half are boys according to the GEMR (2018). Less than 25% are below 12 years old indicating that the problem is predominantly one of adolescents above the minimum age of work of 15 years. Thus 75% of OOSC are of secondary school age (UIS, 2018). This is a radically different view of the OOSC problem than the orthodoxy that emphasises those who do not enrol in grade 1 (Zone 1) and those do not complete most primary grades (Zone 2). Flow analysis immediately draws attention to this counter-intuitive topography of exclusion. Such exclusion is overwhelmingly amongst secondary age children whose learning crisis is determined by lack of access and participation at secondary level.

Various 'silent exclusions' can be mapped. At the simplest level national assessments or high-stakes examinations can be used to estimate what proportion of those enrolled in any given grade are displaying levels of achievement in line with the expectations of national curricula. Performance on standardised international tests can also be used with the obvious disadvantage that these have to be translated into achieve-

ment linked to what is actually taught based on the national curriculum. Averages of performance can be used though. Bands of levels of achievement rather than averages would be better than averages since these can capture improved achievement in ways that reflect the spread of capabilities around a median level. Averages do not do this. Country-level analysis of silent exclusions related to achievement has to absorb system idiosyncrasies, e.g., different ages of entry to school, transition points from lower to upper primary and lower and upper secondary, streaming by capability within schools and tracking into different types of school, and medium of instruction. In the most effective systems the distribution of performance is likely to be less than in the poorest-performing systems. Differentiated teaching at levels tailored to levels of student achievement will be most challenging where the spread in achievement is greatest. Multigrade approaches to learning are well documented (Little, 2006) and many established systems exist (e.g., Activity Based Learning (ABL) in India).

Other forms of systems-level silent exclusion can be assessed and profiled. Most obviously these can be indicated by patterns of attendance, proportions of over-age enrolment, transition point drop-out, and exclusions linked to disability (Ampiah & Adu-Yeboah, 2009; Little, Indika & Rolleston, 2011). Part of the value of doing this lies in the process of estimation which provides opportunities to collate data and sensitise actors to dimensions of exclusion and dynamics that may be largely invisible. This can then be the basis for evidence-based policy dialogue linked to dynamic management of change at system level rather than single-issue interventions focused on improving narrowly defined outcomes like age-specific literacy rates.

Generic Patterns of Participation in LICs and LMICs

Analysis across more than 60 education systems shows that there are five characteristic patterns of enrolment in grades 1 to grade 12 in LICs and LMICs (Lewin, 2015, 2017). The synthesis (Figure 6.5) uses an Index that compares enrolments in each grade with the population in the relevant age group. This is broadly equivalent to an age-specific enrolment rate. The index value is 1 when enrolments match the number in the appropriate age group. The five patterns are (1) convex, (2) highly

convex, (3) linear decline, (4) concave, and (5) linear full enrolment. Lists of countries for each category are available in Lewin (2015, 2017).
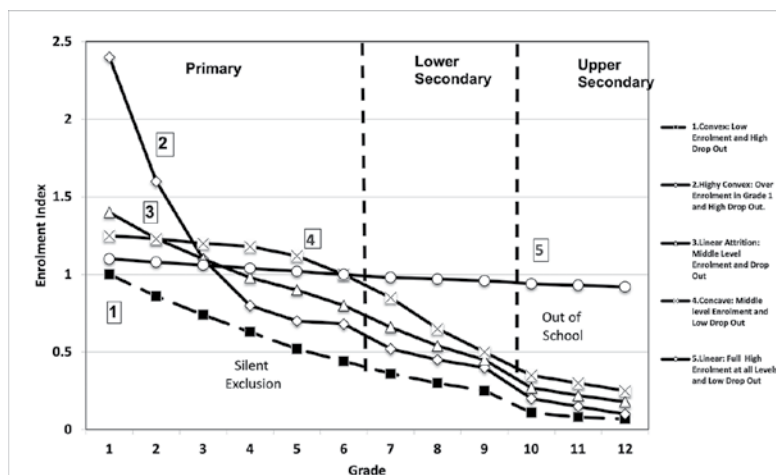
Source: Lewin (2008, 2017, 2021).

Figure 6.5   Types of enrolment by grade in LICs and LMICs

1. Type 1 countries have concave enrolment curves. Intake levels into grade 1 are similar to the number of children in the entry age group. The participation index (number enrolled/number in age group for grade) is close to 1 for grade 1. The tipping point, where there are as many children in the age group than are enrolled in school, is in grade 1 or grade 2. Drop-out starts in grade 1 and results in fewer than 50% completing grade 6. Completion rates may be below 40% at primary and are less than 20% for lower secondary. Development at secondary level is strongly constrained by the output from primary. The priority in these countries is to increase age-related entry and progression rates and reduce drop-out.
2. Type 2 countries have very high rates of over-enrolment in the early grades of primary. Enrolment curves are very concave and tipping points are typically around grades 3. Enrolment in grade 1 may exceed 200% of the number of children in the age group. High drop-out means that less than 70% of the age group complete grade

6 and less than 50% reach grade 9. Over-enrolment arises from many children entering who are over age, and from high rates of repetition. In some countries this pattern has persisted for more than a decade after universal primary education policy has been announced. The implication is that one equilibrium with low enrolment, low drop-out and low completion has been replaced by another with a very high intake, high enrolments and a higher rate of drop-out leading to low completion rates. These countries are often heavily aided.

3. Type 3 includes countries where the intake rate to grade 1 is high but is less than 50% times greater than the number of six-year-olds, and is therefore less than Type 2. Enrolments decline linearly with increasing grade, and the tipping point is around grade 4. No more than 75% of children in an age group reach the end of primary school. There may be serious issues with over-age children and repetition, and with persistent drop-out that accumulates from grade to grade such that fewer than 50% complete lower secondary. Primary completion rates constrain expansion of secondary school.

4. Type 4 include countries that are close to achieving universal completion of grade 6 but have yet to reach more than 50% completing grade 9. Enrolment curves are concave and tipping points are around grade 6 or higher. These countries are more likely to have regularised intake into grade 1 so that all children are within a year of the appropriate age. Most of those who start primary school finish on schedule at the right age. The biggest attrition occurs in lower secondary and less than half of all children succeed in entering upper secondary. These systems are most likely to need support at the post-primary level for curriculum development, quality improvement and enhanced equity as well as investment in infrastructure.

5. Type 5 countries have full enrolment with similar numbers of children enrolled in each grade as there are in the relevant age cohort. Enrolment curves are linear and track the population growth of single age cohorts of children. There is no tipping point. There may also be evidence of demographic transition where the number of children in the single age population declines each year. These systems have achieved universal enrolment up to the end of lower secondary. They are likely to have problems with quality, achievement and equity that would benefit from additional investment.

All systems have quality and achievement issues not evident from simple enrolment flow data. LICs are concentrated in Types 1, 2 and 3. LMICs

are predominantly Type 4 and Type 5 systems. The rates of progress and directions of travel vary. Time series analysis suggests that a proportion of Type 1 LMICs will graduate to become Type 2 or Type 3 within the next decade. It is also probable that Type 1 LICs will become Type 2 systems and some Type 2 will become Type 3. There remains a considerable distance to travel for most LICs and LMICs to become Type 5 full enrolment systems. Until they do, large inequalities are likely to remain. There are also clusters of countries where progress has stalled and the progression through different enrolment types appears to have reached equilibrium states that are difficult to dislodge. In all countries on which there is good time series data the most significant correlates of exclusion, and of learning achievement, are household wealth, followed by location and disability, and then by being a boy or girl.

This analysis shows how impossible it is to reach full enrolment to grade 12 in LICs and LMICs by 2030. For this to happen all children now in grade 1 and 2 would have to remain in school and stay on schedule for the next eight years with no drop-out or repetition. Enrolments in grades 11 and 12 would have to increase by five times or more. Figure 6.4 illustrates the scale of the problem of enrolment increases, teacher supply and school facilities. The CREATE systems analysis highlights how enrolments, and education systems capacity at different levels, would have to change in unprecedented ways to achieve SDG4. Much could be achieved to enhance access to learning through incremental and sustainable system expansion that did not headline unattainable goals but identified achievable goals and timelines that could be resourced.

Patterns of Participation of Boys and Girls

Aggregate enrolment patterns mask differences within populations of children. An illustration of this is to explore how the enrolment of girls and boys varies between countries. Analysis of time series data across 60 countries shows that since 1980 there were dramatic improvements in the ratio of girls to boys enrolled (GEMR, 2018). The detailed patterns are complex but overall in 1990 the Gender Parity Index (GPI) for all developing countries for primary enrolment was 0.87 and for SSA 0.87. By 2015 the value was 0.99 and for SSA 0.94. At secondary level the GPI had reached 0.96 globally and 0.88 in SSA. In all regions girls out-enrolled boys at tertiary level except in South Asia and SSA. The exclusion of boys has become much more visible especially amongst older age groups at higher educational levels (GEMR, 2018). Few

would have predicted that by 2015 girls would out-enrol boys in higher education in Europe, North and South America, and the Caribbean by more 130 to 100, and would be outperforming boys in many areas of the curriculum. Gendered exclusion from school has diminished greatly especially in LMICs but it persists especially for children of the poorest households and in the lowest income countries where boys persist longer in school than girls.

Using a simple parity index[3] indicates the percentage of girls enrolled by grade (Figure 6.6). The results vary widely across groups of countries. Data for the percentage of girls enrolled by grade from around 2015 for SSA produces four clusters of countries with distinctly different patterns of enrolment.
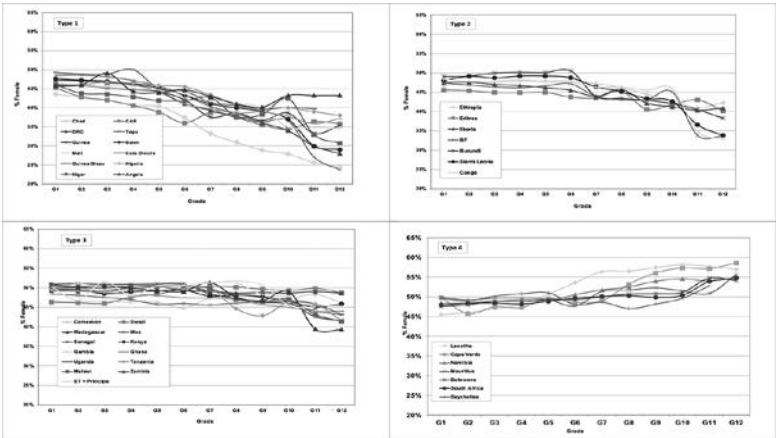
Source: Author's charts derived from UIS data.

Figure 6.6   Enrolment of girls as a percentage of the total in SSA

In the first cluster of countries there are large differences in enrolment in favour of boys throughout these education systems. They also are likely to have a low overall level of participation for both boys and girls. The

---

³   This index simply computes the percentage of girls enrolled of the total number. It therefore does not correct for any imbalances in the population of girls and boys.

priority in these countries is likely to be to invest in interventions that increase participation of both girls and boys to much higher levels since higher participation almost invariably leads to more equal enrolments.

In the second group of countries enrolments of girls are between 45% and 50% of the total through primary grades. Above this grade level there is strong attrition of girls coinciding with exclusions from secondary grades and premature drop-out. Exclusion is often concentrated amongst particular sub-populations, e.g., the poorest households, specific social groups, and those in particular geographic areas whose enrolment rates are likely to be low.

The third group of countries is the largest and includes countries that have equal enrolments of girls and boys up to the end of primary (if equity is defined as a ratio of girls to boys of 48%–52%). At secondary level, girls' participation begins to fall off but generally remains above 45%. This is despite over-age progression, early marriage, underachievement, low returns for household investment, and social prejudices against the education of girls. Which factors are most important are country specific.

The fourth group of countries has close to full enrolment of girls and boys. In these countries there is a tendency for girls to out-enrol boys especially in higher enrolment countries. This may or may not conceal differences within particular groups and regions.

Overall, there are four different patterns of exclusion in LICs and LMICs. These can be described as (1) strong exclusion of girls in all grades; (2), weak exclusion of girls in primary, strong exclusion at secondary; (3) near equity in primary and weak exclusion of girls at secondary; and (4) equity or enrolment of more girls than boys in most grades. The percentage of girls enrolled at each grade level is illustrated with single country cases (Figure 6.7). A full list of countries in each category is available in Lewin (2015).

In Pattern 1, 80% of girls and boys have similar enrolment status to each other at primary level. Only 5% of countries are in Type 1. In Pattern 2, 90% girls and boys have the same participation rates as each other. The problem of more equitable enrolment is concentrated amongst the 10% of children that have different enrolment status suggesting that sharply targeted interventions are much more likely to have an impact on the differences. In Pattern 3 and 4 the great majority of girls and boys (well over 90%) have the same enrolment status. This does not mean that equity is achieved. It does mean that indicators other than enrolment and completion rates are needed to identify, monitor and reduce forms of differential exclusions of girls or boys.
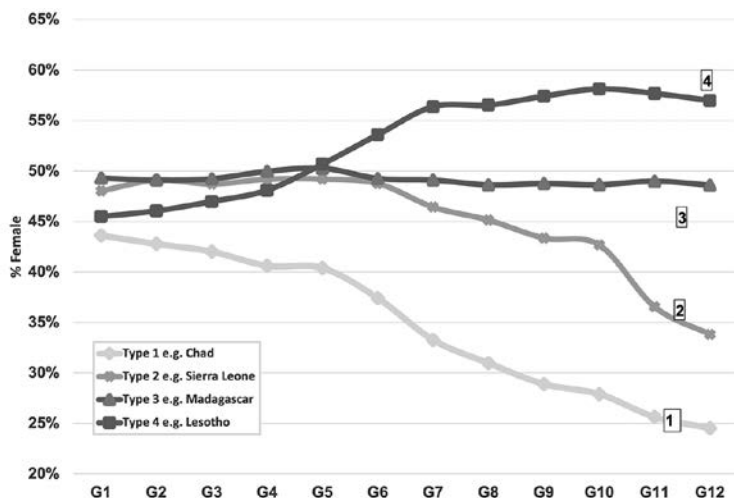
Source: Updated from Lewin (2017).

Figure 6.7   LICs and LMICs classified by percentage of girls by grade

Analysis of the data sets indicates that in LICs and LMICs (1) differences in enrolment of girls and boys tend to diminish as enrolment rates increase and patterns 3 and 4 become the most common; (2) differences in enrolments between boys and girls are larger for secondary schools than for primary; (3) where enrolment rates at secondary are above 50% girls tend to out-enrol boys; (4) in SSA in most countries girls tend to enrol younger and leave school earlier than boys who repeat more often and remain until greater ages. Time sequence data show that most LICs and LMICs have made substantial progress towards equal enrolments and 75% of LICs and LMICs are now either Type 3 or type 4. In contrast, data on inequalities related to wealth shows much greater discrimination in chances of enrolment and less change or consistency in the direction of travel. Systems analysis shows how gender inequities differ radically across countries and within systems at different grade levels.

CONCLUDING REMARKS

This synthesis illustrates the power of the approaches developed by CREATE for the analysis of systems. It has provided an introduction to some of the CREATE conceptual models and to flow analysis using cross-country and sequential cross-sectional data sets. This method of mapping zones of exclusion has proved useful in unpacking key issues, e.g., who fails to enrol at any age, who enrols late, who experiences age grade slippage, how low achievement excludes especially in monograde systems, and how selection examinations exclude those who may already be disadvantaged. The CREATE expanded vision of access and learning and research on political economy and case studies of change provides an entry into system-level discussion of strategy and priorities and the trade-offs that occur in all real-world political systems. Country-analytic reviews and synthetic country research reports provide a mechanism to locate research results in relation to national priorities and highlight system-level interventions which interact and may have positive or negative synergies. All education systems have characteristic arenas which shape how educational services are delivered. These determine who participates, what is available to whom, and how inputs relate to process and outcomes. Mapping these also draws attention to the need to understand that each arena may can be understood with data at different levels of analysis (e.g., the individual, household, community, district authority and schools).

The illustrative system-level data discussed on the evolution of enrolments demonstrate the power of this systems approach to provide critical insights into how participation is changing and why changes may not map well onto SDG4 goals and targets. Exploring the flows quickly reveals that high average enrolment rates at primary level conceal the fact that some countries have massive over-enrolment in low grades and high attrition that persists over time and others do not (Lewin, 2017). The results are a clear signal that education systems have very different trajectories of development that have to be understood, especially where national goals and priorities diverge from global goals. Similar analysis techniques can be used to describe and begin to understand how and why patterns of participation of girls and boys in education systems are changing. Many different patterns of enrolment by grade exist, all located in specific national contexts and system level practices and preferences. Understanding context and history are central to enhancing participation

and learning and rising to the challenges of development (Akyeampong, 2009; Chimombo, 2009; Ahmed & Govinda, 2010; Djangmah, 2011).

CREATE's research portfolio has identified many possible actions that will shape education systems over the next development decade. It developed an architecture of enquiry and related toolkits that could profile the kind of system-level interventions that would make the difference between more uneven and insecure efforts to improve participation and learning, and those interventions that could secure the right to education for all children.

CREATE developed a generic 12-point framework based on its insights and evidence to support new approaches to educational system reform in LICs and LMICs (Box 6.2).

BOX 6.2   CREATE 12-POINT FRAMEWORK FOR SYSTEM DEVELOPMENT

1. Extend early childhood development and pre-school to enhance health, reduce under-nutrition, eliminate stunting and reduce parasitic infections and avoidable causes of disability, and increase school readiness.
2. Ensure that all children enter school at the right age and progress age in grade without slippage.
3. Identify and act on the causes of drop-out on both the supply and demand sides, intervene to enrol students in school below the school leaving age; encourage school-community initiatives to reach and retain all children.
4. Diagnose the silent exclusions that result in enrolment with little learning; track children's progress, manage learning systematically to ensure no child is left behind with clear loci of responsibility.
5. Increase equitable access to secondary and higher education at affordable public and household costs.
6. Promote effective pedagogies suited to teacher capabilities and school and class size and learning conditions.
7. Build enough schools, classrooms and educational spaces with low environmental impact.
8. Provide learning materials that are fit for purpose and invest in appropriate infrastructure for learning.

9. Train, employ and deploy teachers to support learning opportunities that are similar from school to school.
10. Use school-based formative assessment to monitor + improve learning; reform high-stakes exams.
11. Ensure financing for balanced pro-poor educational growth from domestic revenues and strategic aid.
12. Develop indicators of progress to monitor equity and efficiency and widen access to all levels of education.

Source: Lewin (2011a, 2015).

This rubric of thematic priorities across systems was developed with four caveats that provide reminders for future systems analysis. First, initial conditions and baseline data have to be analysed as a precursor to developing plans to improve access and learning opportunities (Lewin, 2015; Little, 2021). Insights into the evolution of participation over the last two decades can give clear indications of the nature of the problems and the likely patterns of causality that continue to deny rights to education to different groups of children. Thus, for example, many systems have pinch points around selection examinations that ration access and exacerbate inequalities, especially those related to household wealth and differences between boys and girls. Most out of school children were in school at some previous point in time, and intervening before drop-out is almost always cheaper than doing so after drop-out. Critically, national systems are very different to each other when data are disaggregated (Lewin, 2007a). Dynamic modelling linked to questions generated by decision makers on understanding and managing flows of learners through education systems provides powerful tools to see beyond the narrative's project-based interventions to identify underlying drivers of change. It is critical to understand why problems that persist have not already been addressed by systems that have experienced low levels of access, participation and achievement and high levels of inequality for decades.

Second, the political economy of reform is such that unless there is sustained local and national political will, availability of adequate resources, and accountability that ensures efficient utilisation, attempts to achieve greater educational participation are likely to prove futile (Little, 2010a, 2010b, 2010c; Obanya, 2011; Ward, 2011). A key difference between low-income countries which have succeeded and countries which have failed to develop their education systems effectively lies in consistent

political will. This has to be accompanied by transition to systems that are endogenously financed and thus nationally owned and sustainable (Al-Samarrai, 2007).

Third, approaches to reform have to recognise that children, and the households of which they are part, exist within a web of relationships that will determine what learning they enjoy, how supply and demand for education interact to generate opportunities for learning that has utility, and what uses young people make of the knowledge and skills they acquire. New approaches must also be specific and targeted where there are barriers and disincentives to go to school and to learn, and where system structures interact with agency to lead to premature exit from schooling or silent exclusions within schools. Households, communities, schools and local education authorities all play a role in shaping opportunities and removing inhibitors to participation and learning and they are part of open systems.

Fourth, the CREATE 12-point framework for system development is not a blueprint but a road map which is designed to chart key dimensions of system development. As a map it is permissive of equifinality (different pathways to the same goals) and multifinality (different goals and overlapping pathways). Both are needed for mobilisation in open systems that have different starting points, resource availability, development aspirations and political preferences. To be useful operationally plans need to build from a bespoke diagnosis grounded in system-level data linked to political economy and recent history. Most importantly system development needs to be endogenously driven so that the theory of change that it related to would be owned and embedded rather than borrowed and gifted by distant stakeholders.

Understanding why education systems do not change is at least as important as identifying the correlates and causes of successful reforms that are sustained over decades. That is why evaluating flows over time and exploring the history of system-level educational reforms is critical to inform future policy and practice. If there is puzzlement that learning 'can get … so bad … even though access to enrolment and schooling completed is expanding rapidly' (Pritchett, 2015) this is not difficult to explain if the patterns of expansion are understood. Poorly articulated expansion, especially that which is exogenously driven, can both degrade quality and selectively change the characteristics of the age cohort reaching the end of an educational cycle with predictable effects on achievement. Game theorists might also note that for many players in the 'game' of educational participation expanded access to learning may

well be more of a zero sum than a win-win game. High-stakes selection will continue to pit candidates against each other with more losers than winners. Expansion at one level will reduce the economic rate of return at that level. Many social scientists have pointed out that the social functions of educations systems are rather more important determinants of system evolution and personal and professional motivation than abstract commitments to national development, global goals that have little domestic ownership or rights-based entitlements. It is commonplace to argue that education systems are more about social mobility and social role allocation than they are about learning for many who take part. Juxtaposing access and learning as being in competition with each other is a false dichotomy since access to education is meaningless without learning that has utility. The better question is what utility does access and learning have for whom?

The recent history of system evolution is a good indicator of the potential of systems for change that is sustainable. One of the best indicators of 'what works' in terms of system development is what worked in the past. This is generally a better guide than speculative projections of what might work in the future (GEEAP, 2020, Lewin, 2020b). In the UK, the most likely weather forecast in the absence of comprehensive data is that tomorrow will be like today. This turns out to be true in general but not of course every day! If the most likely pathways of system development are predicted by answering the system-level evolutionary questions 'what has been happening in the recent past?' and 'what will happen if we allow the system to evolve', then a basis exists for managing change and doing much better than chance and avoiding traps set by supposing anything is possible no matter how improbable. Leadership is of course about much more than working out which direction a system is going and 'walking in front'. Systems analysis can be an invitation to be systematic about diagnosis and methods of managing change for which there is effective demand and finding ways of catalysing change which take systems through tipping points to sustained gains in efficiency and effectiveness.

Radical change is generally accompanied by systemic risk. This is poorly matched to fragile systems with low resilience. Conventional planning places too much value on ambitious, short-term goal achievement and undervalues longer-term change that makes use of evolutionary system drivers. Over-challenging goals set in stone at single points in time are progressively overtaken by events leading to policymaking in Zones of Improbable Progress (ZIPs) (Lewin, 2007b) where few plausible analysts believe the goals can be achieved. This can result in poor

allocative decisions, especially where systems have to be financed with temporary injections of concessionary aid and an unpredictable sequence of performance-related contracts. Pilots do need planes to fly and an assured fuel supply, and education systems need schools with teachers and reliable recurrent finance from domestic revenue (Lewin, 2020a).

There is an opportunity to build on effective practice, discourage unwanted outcomes and nudge incremental improvements. Persistent incrementalism within systems that place a value on sustained change is attractive. Equifinality and multifinality are both essential perspectives on managing change in open systems. The diagnosis of system-level problems using systems analysis tools can provide insight into both probable and possible system futures that are achievable and financially and educationally sustainable. A systems approach to development has to align educational reform with national development priorities, opportunities and preferences. The clear implication is that sustainable educational development will depend on systems located differently in time and space and in different political economies of transition. The international architecture for external assistance and the targets set by SDG4 urgently need reform to meet the different challenges this creates for different systems.

REFERENCES

Ahmed, M. (2011). The Sector-wide Approach in Bangladesh Primary Education: A Critical View. CREATE Pathways to Access Research Monograph No. 57.

Ahmed, M. & Govinda, R. (Ed.). (2010). Universal primary education in South Asia: A right that remains elusive. *Prospects*, 40 (3): 322–335.

Akyeampong, K. (2009). Revisiting free compulsory universal basic education (FCUBE) in Ghana. *Comparative Education*, 45 (2): 175–196.

Alexander, R. (2008). EFA, the Quality Imperative and the Problem of Pedagogy. CREATE Pathways to Access Research Monograph No. 20.

Al-Samarrai, S. (2007). Financing Basic Education in Bangladesh. CREATE Pathways to Access Research Monograph No. 12.

Ampiah, J. G. & Adu-Yeboah, C. (2009). Mapping the incidence of school dropouts: A case study of communities in Northern Ghana. *Comparative Education*, 45 (2): 219–232.

Ananga, E. (2011). Typology of school dropout: The dimensions and dynamics of dropout in Ghana. *International Journal and Educational Development*, 31 (4): 374–381.

Anderson, C. A. & Bowman, M. J. (1965). *Education and Economic Development*. Aldine Publishing.

Beharry, G. (2021). The pathway to progress on SDG 4 requires the global education architecture to focus on foundational learning and to hold our-

selves accountable for achieving it. *International Journal of Educational Development*, 82 (82): 1–7.

Burnett, N. (2019). It's past time to fix the broken international architecture for education. *International Journal of Educational Development*, 68: C 15–19.

Cameron, S. (2011). Whether and where to enroll? Choosing a primary school in the slums of urban Dhaka, Bangladesh. *International Journal and Educational Development*, 31 (4): 357–366.

Chimombo, J. (2009). Changing patterns of access to basic education in Malawi: A story of a mixed bag? *Comparative Education*, 45 (2): 297–312.

Colclough, C. L. & Lewin, K. M. (1990). *Educating All the Children: Meeting the Economic Challenges of the 1990s*. UNICEF.

Collier, P. (2018). The ethical foundations of aid: Two duties of rescue. In C. Brown & R. Eckersly (Eds), *Oxford Handbook of International Political Theory*. Oxford Handbooks Online.

Coombs, P. H. (1968). *The World Education Crisis: A Systems Analysis*. Oxford University Press.

Coombs, P. H. (1985). *The World Crisis in Education: The View From the 1980s.* Oxford University Press.

Djangmah, J. S. (2011). Inequitable Access to Basic education in Ghana: The Way Forward for FCUBE. CREATE Occasional Paper No 2.

Dore, R. P. (1976). *The Diploma Disease: Education, Qualifications and Development*. Unwin Education.

Filmer, D., Rogers, H., Angrist, N. & Sabarwal, S. (2018). Learning Adjusted Years of Schooling (LAYS). World Bank Policy Research Working Paper 8591.

GEEAP. (2020). Cost-effective Approaches to Improve Global Learning. What Does Recent Evidence Tell us are 'Smart Buys' for Improving Learning in lLow- and Middle-Income Countries? Global Education Evidence Advisory Panel. World Bank.

GEMR (Global Education Monitoring Report). (2018). Did You Know There are Just as Many Boys Out of School as Girls? UNESCO.

Grindle, M. S. & Thomas, J. W. (1991). *Public Choices and Policy Change: The Political Economy of Reform in Developing Countries.* John Hopkins University Press.

Halsey, A. H., Floud, J. & Anderson, C. A. (1965). *Education, Economy and Society*. Free Press.

Heyneman, S. P. & Lee, B. (2016) International organizations and the future of education assistance. *International Journal of Educational Development*, 48 (C): 9–22.

Hunt, F. (2008). Dropping Out from School: A Cross Country Review of the Literature. CREATE Pathways to Access Research Monograph No. 16.

Laugharn, P. (2007). Negotiating Education for Many: Enrolment, Dropout and Persistence in the Community Schools of Kolondieba, Mali. CREATE Pathways to Access Research Monograph No. 14.

Lewin, K. M. (1985). Quality in question: A new agenda for curriculum reform in developing countries. *Comparative Education*, 21 (2): 117–133.

Lewin, K. M. (2007a). Improving Access, Equity and Transitions in Education: Creating a Research Agenda. CREATE Pathways to Access Research Monograph No. 1.

Lewin, K. M. (2007b). Long Term Planning for EFA and the MDGs: Modes and Mechanisms. CREATE Pathways to Access Research Monograph No. 7.

Lewin, K. M. (2008). Strategies for Sustainable Financing of Secondary Schooling in Sub Saharan Africa. World Bank.

Lewin, K. M. (2011a). Making Rights Realities: Researching Educational Access, Transitions and Equity. Research Report. http://www.create-rpc.org/pdf_documents/Making-Rights-Realities-Keith-Lewin-September-2011.pdf

Lewin, K. M. (2011b). Target setting and target getting in EFA policy and in EFA practice. *Journal of Education Policy*, 26 (4): 571–587.

Lewin, K. M. (2015). Educational Access, Equity and Development: Planning to Make Rights Realities. Fundamentals of Educational Planning Serial Number 97. International Institute for Educational Planning, UNESCO.

Lewin, K. M. (2017). The Educational Challenges of Transition: Key Issues for Low and Low Middle Income Countries and the GPE Towards 2030. Occasional Paper 2, Global Partnership for Education.

Lewin, K. M. (2020a). Beyond business as usual: Aid and financing education in sub-Saharan Africa *International Journal of Educational Development*, 78 (C): 14.

Lewin, K. M. (2020b). Smart Buys, Great Sales and Special Offers? Commentary on Report of the Global Education Advisory Panel, November 2020. UK Forum for International Education and Training (UKFIET) Blog online. https://www.ukfiet.org/2020/smart-buys-great-sales-and-special-offers-cost-effective-approaches-to-improve-global-learning/

Lewin, K. M. (2021). Unequal access to education: Accounting for change and counting costs. In P. Rose, M. Arnot, R. Jeffery & N. Singal (Eds), *Reforming Education and Challenging Inequalities in Southern Contexts*. Routledge.

Lewin, K. M. & Akyeampong, K. (2009). Education in sub-Saharan Africa: Researching access, transitions and equity. *Comparative Education*, 45 (2):143–151.

Lewin, K. M. & Little, A. W. (2011). Access to education revisited: Equity, drop out and transitions to secondary school in South Asia and Sub-Saharan Africa. *International Journal of Educational Development*, 31 (4): 333–337.

Lewin, K. M. & Sabates, R. (2011). Changing Patterns of Access to Education in Anglophone and Francophone Countries in Sub Saharan Africa: Is Education for All Pro-Poor? CREATE Pathways to Access Research Monograph No. 52.

Lewin, K. M. & Stuart, J. S. (2003). Researching teacher education: New perspectives on practice, performance and policy multi-site teacher education research project (MUSTER). Department for International Development. https://assets.publishing.service.gov.uk/media/57a08d1d40f0b652dd001780/Educationalpaper49a.pdf

Little, A. W. (Ed.). (2006). *Education for All and Multigrade Teaching: Challenges and Opportunities.* Springer.

Little, A. W. (2008). Politics, Policies and Progress. CREATE Pathways to Access Research Monograph No. 13.

Little, A. W. (2010a). The Politics, Policies and Progress of Basic Education in Sri Lanka. CREATE Pathways to Access Research Monograph No. 38.

Little, A. W. (2010b). Access to Basic Education in Ghana: Politics, Policies and Progress. CREATE Pathways to Access Research Monograph No. 42.

Little, A. W. (2010c). Access to Elementary Education in India: Politics, Policies and Progress. CREATE Pathways to Access Research Monograph No. 44.

Little, A .W. (2021). Education for all in India and Sri Lanka: The drivers and interests shaping egalitarian reforms. In P. Rose, M. Arnot, R. Jeffery & N. Singal (Eds), *Reforming Education and Challenging Inequalities in Southern Contexts*. Routledge.

Little, A. W. & Lewin, K. M. (2011). The policies, politics and progress of access to basic education. *Journal of Education Policy*, 26 (4): 477–482.

Little, A. W., Indika, U. & Rolleston, C. (2011). Access, Attendance and Achievement in Rural Schools in Sri Lanka. CREATE Pathways to Access Research Monograph No 73.

LMTF (2015). Access plus Learning. Brookings Institution. https://www.brookings.edu/product/learning-metrics-task-force/

Lockheed, M. E. & Verspoor, A. (1991). *Improving Primary Education in Developing Countries*. Oxford University Press.

Obanya, P. (2011). Politics and the Dilemma of Meaningful Access to Education: The Nigerian Story. CREATE Pathways to Access Research Monograph No. 55.

Oketch, M. & Somerset, A. (2010). Free Primary Education and After in Kenya: Enrolment Impact, Quality Effects, and the Transition to Secondary School. CREATE Pathways to Access Research Monograph No. 37.

Pritchett, L. (2015). Creating Education Systems Coherent for Learning Outcomes. RISE Working Paper Series 15/005. https://doi.org/10.35489/BSG-RISE-WP_2015/005

Spivak, M. (2021). Applying Systems Thinking to Education: The RISE Systems Framework. https://riseprogramme.org/publications/applying-systems-thinking-education-rise-systems-framework

UIS. (2018). One in Five Children, Adolescents and Young People is Out of School. Fact Sheet No. 48, February 2018 UIS/FS/2018/ED/48. UNESCO Institute of Statistics. https://uis.unesco.org/sites/default/files/documents/fs48-one-five-children-adolescents-youth-out-school-2018-en.pdf

UNESCO. (1990). Framework for Action to Meet Basic Learning Needs, World Declaration on Education for All. UNESCO.

UNESCO Institute for Statistics. (n.d.). *UIS Stat.* http://data.uis.unesco.org

Ward, M. (2011). Aid to education: The case of Sarva Shiksha Abhiyan in India and the role of development partners. *Journal of Education Policy*, 26 (4): 543–556.

World Bank. (1980). Education Sector Policy Paper. World Bank.

Zeitlyn, B. & Hossain, A. (2011). Access to Education and Age in Grade in Bangladesh. CREATE Bangladesh Policy Brief No. 5.

7. System coherence for learning: applications of the RISE education systems framework

Michelle Kaffenberger and Marla Spivack

INTRODUCTION

In recent decades, the education systems in most low- and middle-income countries (LMICs) have succeeded in rapidly expanding access to schooling so that, today, most children attend at least some amount of schooling. However, this expansion in access to schooling has not produced commensurate improvements in learning outcomes. A global learning crisis has been widely acknowledged (World Bank, 2018; Angrist et al., 2021; Beatty et al., 2021; UNESCO et al., 2021).

There are many potential drivers of low learning in LMICs. For example, financing may be inadequate (UNESCO, 2015); teaching and learning materials may be lacking (UNESCO, 2016); teachers may be in short supply or poorly prepared (Education Commission, 2019); or schools may be poorly managed (Lemos et al., 2021). These challenges represent proximate determinants of the learning crisis, or the determinants most directly associated with low learning (Pritchett, 2015).

Furthermore, the effects of many proximate determinants of learning vary widely across contexts (Pritchett, 2021). One study finds that the effect of smaller class sizes varied from negative to zero to positive based on the setting (Wößmann & West, 2006). A recent report by the World Bank compared the effectiveness of different categories of learning interventions. Programmes in one category designated as a 'good buy' (structured lesson plans with linked materials, teacher training and monitoring) ranged from the least effective (i.e., a negative impact on learning outcomes) to the second most effective of all included interventions (World Bank, 2020).

A systems thinking approach seeks to understand what system dynamics produced the proximate determinants of low learning. It further helps diagnose the system dynamics that drive the varied effects of proximate determinants in different contexts. Ultimately, systems thinking seeks to inform action to address underlying system dynamics.

The Research on Improving Systems of Education (RISE) programme applies systems thinking and analysis to understand why learning is low and how education systems can shift to improve outcomes. The programme's research agenda is anchored around an education systems framework that specifies the elements, relationships and feedback loops in education systems and the ways these interact to achieve or frustrate children's learning. The framework hypothesises that low learning is the result of education systems that are primarily coherent for schooling rather than for learning and that systems deliver learning when their elements are coherent for learning objectives.

The RISE systems framework helps to understand success, diagnose failure and inform actions that, taking into account the interactions in a system, could bring about change. It can be used to understand education system outcomes at different levels of schooling (e.g., primary, secondary) and different levels of the education system (e.g., municipal, state, national). This chapter will describe the RISE education systems framework and provide two applications of the framework to empirical examples of system coherence and incoherence for learning.

THE RISE EDUCATION SYSTEMS FRAMEWORK

The RISE education systems framework provides scaffolding for considering the key elements, actors and relationships in an education system and the ways these interact to produce a system's outcomes (Pritchett, 2015).[1]

The RISE framework is rooted in the conceptualisation of service delivery systems presented in the 2004 *World development report: Making services work for poor people*. This report frames service delivery systems as a set of interconnected principal–agent relationships, which are referred to as an accountability triangle (World Bank, 2004). This conceptualisation, in turn, was underpinned by the 'strategic tri-

---

[1]   This section draws heavily on Pritchett (2015); a summary of the framework is provided in Spivack (2021).

angle' articulated in Moore's (1995) *Creating public value: Strategic management in government*. The RISE framework adapts the WDR accountability triangle to describe the education sector. It also draws on insights into how states build capability for implementation, in particular the importance of careful identification of a problem and its root causes, before attempting to develop a solution (Andrews, 2017).

Actors and Relationships

Education systems are made up of many actors. The RISE framework summarises these actors as citizens (parents, children, communities, etc.); executive, legislative and fiduciary authorities; education authorities and organisations; and frontline workers (school leaders, headteachers, teachers, etc.). It uses the paradigm of a relationship of accountability with a principal and an agent to describe their interactions (Figure 7.1) (Pritchett, 2015; World Bank, 2004). In its simplest form, this frames the relationship in terms of a principal wanting a task accomplished and engaging with an agent to complete the task. For example, a ministry of education wants children to be taught, so it engages with teachers to teach children.

The RISE systems framework includes four key relationships of accountability between these actors (Figure 7.2) (Pritchett, 2015; World Bank, 2004). First, the 'politics' relationship is the relationship between citizens who are the principals, and the highest executive, legislative and fiduciary authorities of the state (e.g., the president or prime minister's office, the parliament, and the finance ministry), which are the agents. This represents the ways citizens express preferences to political actors and, ideally, hold political actors accountable for their actions. Among citizens, there may be groups or coalitions with varying degrees of influence in the 'politics' relationship. For example, the wealthy or privileged may have more influence than the poor or marginalised.

Second is the 'compact' relationship. In this relationship the highest executive, legislative and fiduciary authorities of the state are the principals, and the education authorities and organisations are the agents. In this relationship, (non-education) authorities, such as the ministry of finance or legislature, interact with education actors such as the ministry of education through actions that can include determining budgets or delegating priorities.

Third, 'management' is the relationship between education authorities and organisations, which here serve as the principals, and frontline
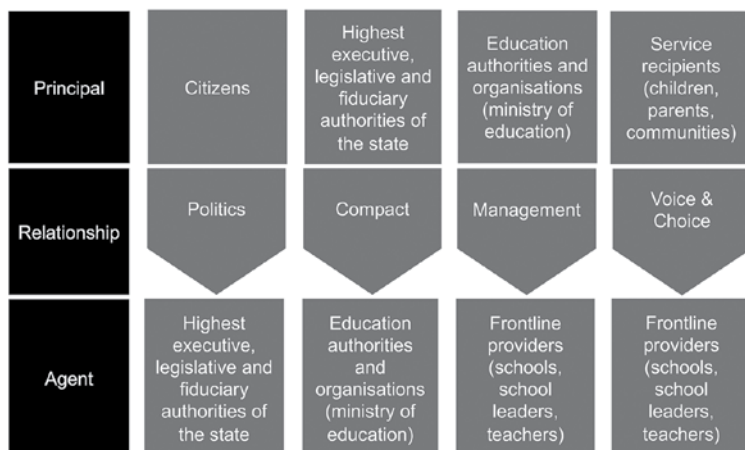
Source: Spivack (2021), adapted from Pritchett (2015).

Figure 7.1   *Four accountability relationships in the education system*

workers, such as school leaders and teachers, which here are the agents. The education authorities include all levels of the education bureaucracy, and the dynamics of this relationship vary based on factors such as the level of (de)centralisation.[2]

Fourth, 'voice and choice' involves the relationship between the recipients of services, including parents, children and communities, who here are the principals, and the frontline workers who provide services, including school leaders and teachers, who here are the agents. As part of this relationship, frontline providers provide instructional services to children.

---

[2]   In many education systems, the 'management' relationship exists within a single organisation – the ministry of education. This would be the case if all or most education functions fall under the remit of a single ministry. In other education systems, the relationship is more complex, with multiple organisations in the 'education authority' role, and each with their own set of frontline workers. For example, in some systems, there are schools that fall under the authority of a ministry of education and other schools that are managed by a religious authority.

Figure 7.2   Relationships and actors in the education system

Design Elements

Actors in an education system interact in many ways. The RISE framework includes five design elements that cut across each relationship and describe the interactions between the actors. These design elements describe the relationship between the principal and agent in terms of what the principal asks the agent to do, how the principal equips the agent to do it, and how the principal monitors and incentivises the agent's performance (Pritchett, 2015).

The first design element is 'delegation', which is what the principal delegates to or expects the agent to do. The second is 'finance', which refers to the resources the principal has allocated to the agent to achieve the assigned task. The third is the 'information' the principal uses to assess the agent's performance. Fourth is 'support', which refers to the preparation and assistance that the principal provides to the agent to complete the task (e.g., teacher training and instructional materials). Fifth is 'motivation', which refers to how the principal motivates the agent, including the ways in which the agent's welfare is contingent on their performance. Motivation can be external (mediated by the principal, i.e., salary) or internal (mediated by the agent, i.e., job satisfaction).

Combining the four key relationships and five design elements produces a 5x4 matrix (Table 7.1) which represents the RISE systems framework and facilitates analysis of the interactions between the actors in the system and how these interactions produce system outcomes.

Embedded in the RISE framework is the assumption that coherence across the relationships of accountability and design elements matters for the outcomes a system produces (Pritchett, 2015). To produce learning, the framework hypothesises that relationships and design elements need to be at least somewhat aligned with learning objectives.

Systems may be incoherent for learning in at least two ways. First, education systems can be coherent for a goal other than learning. In recent decades, many education systems in LMICs have achieved rapid and large-scale increases in school grade attainment by aligning their systems for schooling access (Pritchett, 2013). Some education systems are aligned to identify and select top performers for elite schooling while leaving most children behind (Muralidharan & Singh, 2021). In these situations, the columns and rows of the RISE framework are largely aligned with each other but work towards a goal other than universal learning.

Second, the relationships of accountability (the columns in the framework) and design elements (the rows in the framework) can be incoherent with each other. For example, within a 'compact' relationship, the executive authority may delegate learning improvements but only ask the education authorities (i.e., a ministry of education) for information on enrolment rates and teacher attendance (Example 1 in Table 7.1). In this case, there is incoherence between what is delegated and the information used to evaluate the ministry's performance. This is an example of incoherence within a relationship of accountability between two design elements.

Incoherence can also arise within a design element (a row) between the relationships of accountability (the columns). For example, the education ministry may launch a new initiative aimed at ensuring that all children master foundational skills and then delegate these priorities to teachers and schools. Parents, however, may prioritise their children passing a high-stakes exam and pressure teachers to prioritise test preparation (Example 2 in Table 7.1).

Of course, this is not the only framework for studying education systems, and other useful frameworks are discussed elsewhere in this book. The contributions of the RISE framework are to provide a structure for enquiry into the systemic roots of low learning outcomes and to emphasise the role of coherence among relationships in driving system

Table 7.1   *The 5x4 education systems framework (five design elements and four relationships of accountability)*

Principal–agent relationships				
Five elements of each relationship	**Politics** Citizens to the highest authorities of the state	**Compact** Highest authorities of the state to education authorities	**Management** Education authorities to frontline providers (schools, school leaders, and teachers)	**Voice & Choice** Service recipients (parents/children) to frontline providers (schools, school leaders, and teachers)
Delegation: What the principal wants the agent to do		Example 1. (a) Executive authority delegates learning improvements	Example 2. (a) Education ministry launches new foundational skills learning initiative	Example 2. (b) Parents prefer and pressure schools and teachers to prioritise preparation for high-stakes school leaving exams
Finance: The resources the principal has allocated to the agent to achieve assigned task				
Information: How the principal assess the agent's performance		Example 1. (b) Despite delegating learning improvements, the executive authority only monitors information on enrolment rates and teacher attendance		
Support: Preparation and assistance that the principal provides to the agent to complete the task				
Motivation: How the principal motivates the agent, including the ways in which the agent's welfare is contingent on their performance against objectives				

Source: Adapted from Pritchett (2015), with examples from the authors.

outcomes. Applying this framework to examine education systems can help identify the incoherence hindering progress.

DELEGATION OF LEARNING GOALS AND SYSTEM ALIGNMENT FOR LEARNING THROUGH THE LENS OF A SYSTEMS FRAMEWORK: THE CASE OF SOBRAL, BRAZIL

A small but growing body of evidence suggests that a key to achieving large improvements in learning is the clear delegation of explicit learning goals by the leadership in an education system, establishing a common purpose and driving other elements of the system to align around the learning goals (Kaffenberger, 2021; London, 2021). The RISE systems framework can be applied to better understand and analyse the sources and processes of such success.

Through the lens of the RISE framework, this section analyses the experience of the municipality of Sobral, Brazil, which has achieved transformative improvements in learning outcomes in recent years.[3] A theme in the Sobral experience is the commitment and dedication of municipal leadership to explicit learning goals and the clear communication and delegation of those goals to the rest of the system. This established a common purpose and collective responsibility for achieving goals, enabling many other elements of the system, including 'information', 'motivation', 'support' and 'finance', to align around the goals.[4]

In just 12 years, Sobral rose from being the 1,366th ranked municipality in Brazil for learning outcomes to being the top performer in Brazil's national basic education assessment (Crouch, 2020). This occurred despite high levels of poverty: in 2017, its test scores were 80% higher than would be expected for its level of education expenditure relative to other Brazilian municipalities.

A key driver of Sobral's learning gains was the clear delegation of explicit learning goals by Sobral's mayor (Loureiro & Cruz, 2020; Crouch, 2020) and subsequent collective commitment to the goals (McNaught, 2022). In 2000–2001, an independent learning assessment

---

[3]   This section draws on case studies of the Sobral experience by Loureiro and Cruz (2020) and Crouch (2020).

[4]   The example of Sobral also shows how the RISE framework can be applied at different levels of the system, including the national, regional or municipal levels.

conducted by the municipality revealed that 40% of primary school students could not read (Loureiro & Cruz, 2020). In response to these findings and others, Sobral's mayor established seven education goals, the top two priorities of which were achieving universal literacy in the first two years of primary school and remediating children in higher grades who could not yet read (Becskehazy & Louzano, 2019). These goals had a slogan – 'Alphabetization (literacy) at the Right Age' – and significant collective responsibility was fostered in support of the goals. In the context of the RISE systems framework, this represents 'delegation' in the 'compact' relationship in which the executive authority (in this case, the municipality's mayor) delegated goals to the education actors in the system. The Secretariat of Education – that is, the education authority in the municipality – then delegated and supported schools and teachers to achieve these goals through the 'management' relationship.

This delegation led to a series of policies and reforms that were coherent with each other and coherent with the delegated learning goals. According to Loureiro and Cruz, Sobral's success was because of 'its ability to converge the whole education system toward learning' (2020, p. 13), with sustained political leadership being an essential condition underlying the other efforts. The efforts involved reforms to curriculum, pedagogy, training and professional development for teachers ('support'); new student assessments used for tracking progress and informing adjustments to classroom instruction ('information'); new incentives and recognition for teachers tied to performance on the learning goals ('motivation'); and increased funding and financial autonomy ('finance') (Table 7.2).

To support teachers and schools in achieving the learning goals, the secretariat of education set clear, sequenced learning objectives, establishing expectations for students at each learning level and grade (Loureiro & Cruz, 2020; Crouch, 2020). A sequenced curriculum was then developed along with structured teaching and learning materials and student assessments, all of which were aligned with the learning objectives. Teachers participated in initial training on the curriculum, learning objectives, pedagogical practices, and materials for use in the classroom, as well as subsequent monthly in-service training. Regular classroom observations from schools' pedagogical coordinators and from secretariat staff provided regular feedback and tailored support for teachers. Secretariat staff visited schools monthly to provide support to coordinators and teachers (Loureiro & Cruz, 2020).

New sources of information were introduced into the system to track and support progress on the delegated learning goals, with 'information about learning outcomes extensively used to guide the education strategy at the municipal, school and classroom levels' (Loureiro & Cruz, 2020, p. 18). Learning assessments were conducted twice per year, with midterm results used to inform course correction and end-of-year results to inform strategies for the following year. Using information on learning was a priority of education leadership – the secretariat dedicated one-third of their time and effort to student assessments, including designing, implementing and analysing assessment results before using these results to provide feedback and guidance on progress to schools (Loureiro & Cruz, 2020). In the classroom, continuous assessment was part of the new, structured pedagogical approach, and teachers were trained and supported in using these assessments to adjust their instruction (Crouch, 2020).

Teachers and other education actors were provided with new incentives to motivate a focus on the delegated learning goals. Financial incentives were established for teachers, pedagogical coordinators, and school principals when the schools achieved annual learning goals, and the teachers could receive bonuses if their class performed well (Loureiro & Cruz, 2020). Non-monetary incentives were also provided, including special honours and public recognition events for high-performing teachers (Crouch, 2020).

Finance reforms also supported the achievement of the learning goals. Around the same time as the reforms, the federal government in Brazil began pooling education resources from the federal, state and municipal governments and redistributing them based on student enrolment, reducing inequality in education financing and increasing per-pupil financing, particularly in poor municipalities, including Sobral (Loureiro & Cruz, 2020). Within the municipality, Sobral undertook a major transition from politically appointed school principals to meritocratically selected principals who were chosen for their technical and pedagogical skills. With skilled leadership in place, Sobral devolved financial autonomy to schools, which came with two main effects. First, the schools had both more financial independence and more responsibility for achieving results through results-based accountability. Second, the role of the secretariat was transformed from a primarily administrative role to a technical one, including providing pedagogical and assessment support to schools (Loureiro & Cruz, 2020).w

Table 7.2   *System reforms in Sobral, Brazil, created coherence for learning across the 'compact' and 'management' relationships and all five design elements*

Principal–agent relationships of accountability				
Five design elements	**Politics**	**Compact**	**Management**	**Voice & Choice**
Delegation		Mayor delegates explicit learning goals, including universal literacy in first two years of primary and remediation for children in older grades, with the slogan of 'Alphabetisation (literacy) at the Right Age'	Secretariat of education delegates goals to schools and teachers and brings other system elements in line with the delegated goals	Parents expressed initial resistance to the reform, but regular dialogue from the mayor and secretariat increased support. Parents were encouraged to reinforce learning goals and ensure that their children attend school more
Finance		Federal education funding increased for poor municipalities, including Sobral	Financial autonomy devolved to school level, giving more financial independence and responsibility for results	
Support			Teachers provided with sequenced learning objectives, structured teaching and learning materials, training and professional development and ongoing feedback and support through classroom observations, all aligned with learning goals	
Inform-ation	Inform-ation on low learning from new assess-ments was shared publicly by the mayor to increase citizen buy-in for improving learning		Use of information on learning a top priority for education leadership, with one-third of the time and effort dedicated to this Twice-yearly assessments used by education leadership to measure progress and inform course correction and strategy Teachers supported the use of continuous assessment in the classroom for regular feedback on student progress and to inform adjustments to instruction	
Motivation			Financial incentives for teachers, in-school pedagogical coordinators and principals for achieving learning goals Public recognition events for high-performing teachers	

Source: Authors' analysis, drawing on Loureiro and Cruz (2020) and Crouch (2020).

Although many education policies and projects aim to make changes to one cell of the RISE systems framework, such as increasing budget outlays (in the 'finance'/'compact' cell) or implementing a teacher training programme (in the 'support'/'management' cell), the Sobral experience stands out for including a coherent set of reforms encompassing many cells of the framework. This integrated, system-wide approach – combined with political commitment to learning and the common purpose and collective responsibility for results at all levels of the system – produced a system shift with large improvements in learning outcomes.

TEACHER CAREER STRUCTURES AND COMPENSATION THROUGH THE LENS OF A SYSTEMS FRAMEWORK: THE CASE OF INDONESIA

The structure of teacher careers – how they are recruited, selected, retained and motivated and their professional norms – all bear critically on their performance and, thus, on their interactions with students. This section describes teacher reform in Indonesia through the lens of the RISE framework, illustrating how the framework can help understand success and diagnose failure in reform efforts.

The structure of teacher recruitment in Indonesia originates in the rapid expansion of the Indonesian schooling system in the late 1970s as part of the Suharto government's National Development Strategy (Huang et al., 2020; World Bank, 1990). The rapid expansion necessitated significant growth in the teacher workforce, which prioritised mass hiring to fill positions, with less emphasis given to ensuring the recruitment of quality candidates and providing them with adequate preparation for the classroom (Huang et al., 2020).

Beginning in the early 1990s, there was growing recognition among international advisors and education officials within the ministry of education that the system was failing to deliver adequate learning and that poor teaching was hindering outcomes (World Bank, 1989, 2013). At the same time, teachers were among the most respected members of many communities and an important political constituency (World Bank, 2013). Teacher groups argued that the income levels and professional status of teachers were key constraints to progress and that teachers needed both better pay and professional status on par with doctors and lawyers. With these factors in mind, a consensus emerged among the various relevant ministries (including education, finance, and planning),

political parties, legislature and teacher groups that a reform effort aimed at improving teachers' performance and rewards could be a viable path forward (World Bank, 2013).

A reform package with three main components was developed with the intent of overhauling the teacher career structure, 're-professionalising' teachers, improving equity in the geographical distribution of teachers and increasing motivation and performance. First, the package sought to improve teacher quality by increasing support for teachers with training and certification. As initially proposed, certification was meant to include an external evaluation of teacher's pedagogical knowledge along with a year of further training and assessment for teachers who failed certification (World Bank, 2013; Ree et al., 2018). Second, the reform sought to increase teacher motivation by tying salary increases to training and certification. Most civil service teachers would qualify for a 100% salary increase if they successfully completed the certification process.[5] Third, the reform provided bonuses for teachers who accepted posts in marginalised areas (World Bank, 2013).

The reform was deployed in the form of a new teacher law (referred to as the '2005 Teacher Law') which was adopted by the legislature and implemented by the Indonesian Ministry of Education. It was primarily financed through a contemporaneous constitutional amendment mandating that 20% of government spending go to education.

Despite intentions, pressure from teachers' associations throughout the policymaking and implementation process diluted the reform, producing an enacted reform that differed substantially from the initial design (World Bank, 2013; Ree et al., 2018). The teachers' groups successfully lobbied to eliminate funding for external teacher evaluations as part of the teacher certification process. External evaluations were replaced with a requirement to submit a portfolio of teaching materials for review (World Bank, 2013). In practice, portfolio reviews became largely pro forma, with most teachers passing. Those who did not could complete a two-week course and take a test, which nearly all candidates passed to get certification (World Bank, 2013; Ree et al., 2018). As a result, the law effectively provided for a nearly universal doubling of civil servant

---

[5]   Technically, the salary increases were only available to teachers with a four-year degree or a sufficiently high civil service ranking. However, most teachers without a four-year degree were administratively given a high enough civil service rank to qualify, making the salary increase effectively available to all teachers (World Bank, 2013).

teacher salaries, with limited or no requirements aimed at raising teacher qualifications (Ree et al., 2018). Bonuses for working in marginalised areas were left in place.

An evaluation of the reform found that despite achieving many of the intended intermediate effects (teachers were more likely to have obtained certification, were happier with their jobs, and were less likely to have a second job), the reform had no effect on teachers' attendance, their subject knowledge or on student learning outcomes (Ree et al., 2018).

This case illustrates two points about coherence (summarised in Table 7.3). First, incoherence can emerge in delegation by citizens and by government authorities (i.e., incoherence within the 'delegation' row, here between the 'politics' column and the 'compact' column). Effort from government officials for reforms to increase teacher pay and improve motivation and support (through certification and training) with the goal of ultimately improving quality was met with resistance from teachers' groups (part of the citizenry) who opposed the motivation and support components and were interested only in the pay increases.

Second, a change to just one element of the teacher's career – 'finance' – had a limited effect on outcomes because it was not paired with reforms to other system elements. It did not change what teachers were delegated, supported, monitored or motivated to do. The final, watered-down version of the certification requirement was so weak that subsequent evaluations found no difference in the performance or knowledge between certified and uncertified teachers (World Bank, 2013), and the universal salary increase did not change incentives or induce greater effort (Ree et al., 2018).

The RISE framework helps understand why teacher reform in Indonesia did not produce learning gains. Because only one system element was adjusted – 'finance' – without (de facto) reform to related elements, teachers did not need to change their practices to benefit from the salary increase. A description of a coherent teacher career policy that attracts, retains and motivates effective teachers is proposed in Hwa and Pritchett (2021); they emphasise that the design elements of a teacher career path must be coherent across the different stages of teacher careers, from pre-service, to novice, to experienced, to veteran teachers, something that the Indonesian reform did not address in its standardised treatment of nearly all civil servant teachers.

Table 7.3   *The 2005 Teacher Reform in Indonesia suffered from incoherence within the 'delegation' row and within the 'management' column*

	Principal–agent relationships of accountability			
Five design elements	**Politics**	**Compact**	**Management**	**Voice & Choice**
Delegation	Teachers' groups argue that higher salaries and professional status will improve performance Pressure from teachers' groups to dilute aspects of the law, in particular, the teacher certification process	Intended reform: delegation from legislative authorities to adopt pay raises for certified teachers to improve learning. Enacted reform: legal provisions on teacher certification significantly diluted, producing a de facto universal salary increase	Intended reform: delegation of quality improvement for teaching through merit-based certification process. Enacted reform: merit-based components replaced with superficial effectively universal certification process	
Finance		Additional financial resources needed for salary increases financed by a constitutional amendment passed around the same time as mandating 20% of government spending go to education	Intended reform: finance provided to raise salaries for teachers who pass external evaluation for merit-based certification. Enacted reform: finance provided to raise salaries for teachers who submit a portfolio and/or complete a two-week course	
Support			Intended reform: comprehensive support and training to teachers who do not pass the certification process. Enacted reform: completion of a two-week course allows nearly automatic certification	
Information			Intended reform: rigorous external evaluation to verify quality of teacher pedagogical knowledge. Enacted reform: teacher quality superficially verified through portfolio review or a two-week course	
Motivation			Intended reform: salary increase for teachers who pass rigorous certification process. Enacted reform: a de facto nearly universal salary increase not contingent on performance	

Source: Authors' analysis, drawing on World Bank (2013) and Ree et al. (2018).

CONCLUSION

In many LMICs the learning crisis is severe. The RISE systems framework is a tool for describing the complex dynamics of a system, including the design elements, relationships of accountability and feedback loops that drive system outcomes. By going beyond the proximate causes of low learning, it applies systems thinking to describe the fundamental drivers of success and failure and to identify constraints to progress in a particular context.

The RISE framework is useful for understanding successful efforts to improve learning outcomes. In the case of Sobral, Brazil, the framework provided a structured way to describe the system actors playing a role in the successful reforms, the interactions between those actors and the system elements that changed as a part of the reform.

The framework is also useful for diagnosing the reasons why reform efforts do not produce the intended or desired outcomes. In the Indonesia case, analysis of the reform through the lens of the framework showed that by only changing one system element ('finance') without related changes to other elements (such as 'delegation', 'motivation' or 'support'), the reform did not sufficiently change the dynamics in the system and bring about the intended learning improvements. Lessons from such retrospective analyses can inform future policy reforms.

The RISE framework can also be applied prospectively to inform action. It can be used to diagnose existing incoherence in an education system and inform the needed actions to improve outcomes. Atuhurra and Kaffenberger (2022) apply the framework to identify incoherence in what different education authorities, including curriculum agencies and exam agencies, expect of teachers in Uganda and Tanzania, hence informing the possible actions to improve alignment. The framework has also been used to diagnose existing incoherence in the education system in Ghana as input to government planning and reform processes.[6] In Ghana, workshops and interviews with the government and other actors were used to diagnose critical areas of incoherence that need to be alleviated to improve learning.

There is no quantified level of coherence that is considered 'sufficient' for learning in an absolute sense. In the Sobral case, the coherent set of reforms produced changes in nine cells of the framework, all

---

6   See https://epg.org.uk/portfolio/ghana-accountability-for-learning-framework/

of which worked collectively towards a common purpose. Atuhurra and Kaffenberger's (2022) application of the framework suggests that improved coherence in two cells of the 'management' column could potentially improve learning in Tanzania and Uganda. Rather than setting a target level of 'sufficient' coherence, the RISE framework can be used as a qualitative tool to improve understanding of system dynamics, identify areas of incoherence and determine in which areas changes are most critical for aligning the system for learning.

Applications of the RISE framework reveal additional complexities. The case studies in this chapter illustrate the role that agents – not just principals – play in shaping priorities and outcomes in an education system. The Indonesian Ministry of Education (which is an agent in the 'compact' relationship) influenced the formation of the reform package enacted by the legislature. In Sobral, schools (which are agents in the 'management' relationship) received more autonomy in financial decision making from the secretariat of education but only after explicit learning goals had been set and qualified leadership put in place.

All education actors, from policymakers to mid-level bureaucrats to teachers at the frontline, are embedded in a system that facilitates and constrains their possible actions. Although it is not possible to reform all components of an education system at once, considering the system-level constraints and incoherence can identify the most promising and feasible pathways to improvement. As the Sobral case demonstrates, not every 'cell' in the RISE systems framework must experience reform to improve learning outcomes. However, in Sobral, enough cells underwent a coherent set of reforms to realign the system for learning. Using the RISE framework to adopt a systems lens can help identify which constraints pose the most critical barriers and must be alleviated to enable change.

REFERENCES

Andrews, M. (2017). *Building state capability: Evidence, analysis, action* (1st edn). Oxford University Press.

Angrist, N., Djankov, S., Goldberg, P. K. & Patrinos, H. A. (2021). Measuring human capital using global learning data. *Nature*, *592*(7854), 403–408.

Atuhurra, J. & Kaffenberger, M. (2022). Measuring education system coherence: Alignment of curriculum standards, examinations, and teacher instruction in Tanzania and Uganda. *International Journal of Educational Development*, *92*(102598). https://doi.org/10.1016/j.ijedudev.2022.102598

Beatty, A., Berkhout, E., Bima, L., Coen, T., Pradhan, M. & Suryadarma, D. (2021). Schooling progress, learning reversal: Indonesia's learning profiles

between 2000 and 2014. *International Journal of Educational Development*, *85*(102436).

Becskehazy, I. & Louzano, P. (2019). *Sobral: A case of educational success in the semiarid northeast.* ComCiencia.

Crouch, L. (2020). *Systems implications for core instructional support lessons from Sobral (Brazil), Puebla (Mexico), and Kenya* (RISE Insight Note). RISE Programme.

Education Commission. (2019). *Transforming the education workforce: Learning teams for a learning generation*. Education Commission.

Huang, A. R., Revina, S., Fillaili, R. & Akhamdi. (2020). *The struggle to recruit good teachers in Indonesia: Institutional and social dysfunctions* (RISE Working Paper 20/041). RISE Programme.

Hwa, Y.-Y. & Pritchett, L. (2021). *Teacher careers in education systems coherent for learning: Choose and curate toward commitment to capable and committed teachers (5Cs)* (RISE Primer). RISE Programme.

Kaffenberger, M. (2021). *Aligning education systems for learning: How systems shift.* RISE programme blog. https://riseprogramme.org/blog/aligning-education-systems-for-learning-how-systems-shift

Lemos, R., Muralidharan, K. & Scur, D. (2021). *Personnel management and school productivity: Evidence from India* (RISE Working Paper 21/063). RISE Programme.

London, J. (2021). *Outlier Vietnam and the problem of embeddedness: Contributions to the political economy of learning* (RISE Working Paper 21/062). RISE Programme.

Loureiro, A. & Cruz, L. (2020). *Achieving world-class education in adverse socioeconomic conditions: The case of Sobral in Brazil*. World Bank.

McNaught, T. (2022). *A problem-driven approach to education reform: The story of Sobral in Brazil* (RISE Insight 2022/039). RISE Programme. https://doi.org/10.35489/BSG-RISE-RI_2022/039

Moore, M. (1995). *Creating public value: Strategic management in government.* Harvard University Press.

Muralidharan, K. & Singh, A. (2021). India's new national education policy: Evidence and challenges. *Science*, *372*(6537), 36–38.

Pritchett, L. (2013). *The rebirth of education: Schooling ain't learning*. Brookings Institution Press.

Pritchett, L. (2015). *Creating education systems coherent for learning outcomes* (RISE Working Paper 15/005). RISE Programme.

Pritchett, L. (2021). *Let's take the con out of randomized control trials in development: The puzzles and paradoxes of external validity, empirically illustrated* (Center for International Development Faculty Working Paper No. 399). Center for International Development.

Ree, J. de, Muralidharan, K., Pradhan, M. & Rogers, H. (2018). Double for nothing? Experimental evidence on an unconditional teacher salary increase in Indonesia. *The Quarterly Journal of Economics*, *133*(2), 993–1039.

Spivack, M. (2021). *Applying systems thinking to education: The RISE systems framework* (RISE Insight Note). RISE Programme.

UNESCO. (2015). *Pricing the right to education: The cost of reaching new targets by 2030* (Education for All Global Monitoring Report, Policy Paper 18). UNESCO.

UNESCO. (2016). *Every child should have a textbook* (Global Education Monitoring Report, Policy Paper 23). UNESCO.

UNESCO, UNICEF & World Bank. (2021). *The state of the global education crisis: A path to recovery* (A joint UNESCO, UNICEF, and World Bank report). UNESCO, UNICEF and World Bank.

World Bank. (1989). *Indonesia – Basic education study* (World Bank, Report no. 7841). World Bank.

World Bank. (1990). *Indonesia – Strategy for a sustained reduction in poverty* (World Bank, Report no. 10009). World Bank.

World Bank. (2004). *World development report 2004: Making Services work for poor people*. World Bank.

World Bank. (2013). *Teacher reform in Indonesia: The role of politics and evidence in policy making* (World Bank Working Paper 75532). World Bank.

World Bank. (2018). *World development report 2018: Learning to realise education's promise.* World Bank.

World Bank. (2020). *Cost-effective approaches to improving global learning.* World Bank.

Wößmann, L. & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, *50*(3), 695–736.

PART III

Applying systems approaches in practice

8. Adapting inclusive systems development (ISD) to vocational education and training (VET) and skills development

Mike Klassen, Sandra Rothboeck and Ailsa Buckley

INTRODUCTION

Over the years, vocational education and training (VET) reforms have made progress to increase access, relevance and quality. However, many countries still struggle to provide an enabling environment that can allow people to obtain the relevant skills to become and remain employable and productively engaged. Many reforms seem not to have had the expected impact on learning and labour market outcomes (Allais, 2020). Although systems thinking has been successfully introduced in general education, it has not been as widely applied to VET. Systems thinking can be used to obtain a deeper understanding of the current situation while taking advantage of the inherent capabilities of stakeholders to come together and solve increasingly complex problems. In some cases, systems thinking has been used to involve VET stakeholders in decision making (CEMETS, 2021; European Training Foundation, 2021); however, many VET reform projects have only done so on an ad hoc basis, failing to engage stakeholders at a deeper and more systematic level.

In this chapter, we argue that a narrow framing of the VET 'system' causes fragmentation that leads to the failure of reform efforts. By system, we mean a bounded set of interacting elements (including actors) and their relationships with a common purpose. Conventional VET development projects hypothesise that the government is the fundamental leverage point for change; thus, this defines the system as government-led

coordination of publicly supported VET and labour market insertion programmes and offerings. Although the provision of an enabling environment and normative principles are clearly the responsibility of the state, well-functioning VET systems operate in close connection with other subsystems (e.g., labour market institutions, general education and economic systems, including businesses and industry associations). Therefore, we draw on the skill ecosystems approach (Buchanan et al., 2017) to define a VET ecosystem with wider boundaries that includes a mix of public and private providers, industry and professional associations, employers of various types and sizes, actors from other levels of education and the learners themselves. Here, the term ecosystem implies interconnections and interdependence between education, social and economic policies and the respective legal and regulatory environments, industry structure and firm-level strategies.

Building on the ideas advocated by skill ecosystems researchers and policymakers, we promote an inclusive systems development (ISD) approach that recognises that complex systems are nonlinear, constantly shifting and changing. An ISD approach identifies and defines the boundaries of complex and overlapping ecosystems and subsystems to better understand their interrelatedness; it involves an analysis of system elements and the interactions and feedback loops between them to reveal bottlenecks and surface root causes. The analysis then feeds into the design of inclusive solutions that alleviate systemic barriers to widespread quality VET services in order to deliver benefits to marginalised actors. By mapping and analysing the key actors, capacities and roles in an ecosystem, it is possible to understand roles, responsibilities, dysfunctional relationships, power dynamics and conflicting interests, actively unlocking the motivations and capabilities of actors to contribute to overall coherent system change. Doing so requires fostering multiple parallel partnerships with different system actors and strengthening their ability to (a) shape a joint understanding and vision for their ecosystem; (b) build trust and collaboration; and (c) assume new functions and behaviours with the appropriate organisational capacities and resources to fulfil them.

The main contribution of this chapter is to explain and demonstrate the advantages of applying an ISD approach to VET projects compared with a more conventional project design, reflecting on the implications this has for design and implementation. The chapter is structured as follows: first, we summarise the key issues from the literature on VET in lower-income countries, especially in Africa, and, from a systems perspective, highlight

promising research approaches that explain failed reforms. Second, we introduce ISD and explain how it diverges from conventional VET project approaches. Third, we show how ISD's four key principles draw on and extend market systems development (MSD) processes, instruments and tools developed over time in private sector development and agriculture projects. Finally, we show an example of a project that has actively employed ISD in its design and implementation.

LITERATURE REVIEW

Mainstream policy discourse has historically been sceptical of the benefits of investment in VET in lower-income countries; however, this view has been changing. The negative view of VET is rooted in a linear narrative that VET provides poor rates of return, which is a conclusion derived from problematic studies from the 1980s (Psacharopoulos, 1981, 1985) 'which became donor orthodoxy by the 1990s' (McGrath et al., 2020b, p. 5). The research claiming poor returns was later challenged for its poor application to many countries in Africa, given the differences in economic structures in those contexts (Bennell, 1996), yet it still led to active discouragement of investments in VET for decades.

Mass expansion of education has had the unintended consequence of positioning VET as a lower status or 'remedial' option, particularly as the number of graduates from general and higher education has greatly outstripped the number of available jobs in many countries. Allais (2020, pp. 13–14) explains how World Bank thinking still rests on the problematic assumption that improving VET will drive industrialisation, rather than framing VET and skill formation as embedded within wider social, economic and political systems. This signals the value of systems thinking in understanding VET.

Unfortunately, recent World Bank working papers (Blattman & Ralston, 2015; McKenzie, 2017) have reinforced a narrow negative view of VET, claiming that skills programmes show very limited positive effects and limited value for money. As McGrath et al. (2020b) argue, World Bank studies have shown a problematic focus on 'externally funded, short-duration vocational programmes in settings where there is an unemployment crisis; not on mainstream VET programmes' (p. 6). By trying to isolate a single intervention and its immediate outcomes, they miss out on understanding VET as ecosystems and the complex set of relationships between learners, schools, employers and government

policy. This myopic view fails to understand the underlying causes of problems.

Fragmented policymaking based on research that does not consider a systems perspective has serious consequences when attempting to apply it in practice. In Ethiopia, like in many other countries, rapid educational expansion without sufficient economic development has led to a focus on education at the expense of deeper structural changes to the economy and, ultimately, a loss of human resources to migration, unemployment and underemployment (Rekiso, 2019). In Uganda, 'policymakers pay lip service to the importance of VET but in practice have reinforced colonial attitudes that academic education is all that matters' (McGrath et al., 2020b, p. 12), leading to mass youth employment and further retrenchment of VET's lower status (Okumu & Bbaale, 2019; Openjuru, 2010).

McGrath et al. (2020a) identify several strands of VET research that promise to overcome these shortfalls. Macro perspectives on policy, systems and institutions emphasise the elements of a systems thinking approach: '[We] need to look into system dynamics for the obstacles and opportunities that will shape the likely success of innovations designed to make VET more inclusive and sustainable' (McGrath et al., 2020b, p. 9). Allais (2020) reinforces the perspective that government policy, industrial strategy and VET systems reform must be linked up: 'High-quality small-scale interventions that are embedded in industrial policy are the most likely to succeed' (p. 14). Growing research on vocational knowledge goes beyond 'crude technical approaches to what skills appear to be needed at the surface level, and to consider what knowledge, as well as skills, is required for transformative VET' (McGrath et al., 2020b, pp. 9–10). There is also an expanding body of research using the critical capabilities approach (McGrath et al., 2020a; Powell & McGrath, 2019), which centres on the aspirations of learners, challenging narrow human capital theory assumptions to take a broader view of how VET can lead to human flourishing, individual empowerment and agency. Swisscontact promotes a similar perspective where skills development (SD) builds the long-term employability and well-being of learners; this combines workplace-related competence (knowledge, skills and attitude) to flexibly adapt to fast-changing environments with lifelong learning and responsible citizenship capabilities (Swisscontact, 2021; Weyer & Kehl, 2015).

These latter approaches from the VET research literature align with skill ecosystems, which is an influential systems thinking framework focusing on the 'content' of skills and 'context' in which skills are used

(Finegold, 1999). The skill ecosystems framework foregrounds the demand side of where and how skills are used in work contexts. The framework has been most actively applied by policymakers and applied researchers in Australia, Scotland/the UK and the US to stimulate links among enterprises, professional bodies, trade unions and education system actors (Buchanan et al., 2017). However, in practice, it has run into significant implementation challenges: 'employer engagement has emerged as arguably the most problematic aspect of skill ecosystem reform' (Buchanan et al., 2017, p. 453). There is a need to bring together a broad range of actors and focus on building networks with genuine ownership from employers. These implementation issues are crucially important for development programmes that seek to support systemic change in VET ecosystems, which is the focus of this chapter.

ISD FOR SKILLS DEVELOPMENT (ISD-SD) IN CONTRAST TO CONVENTIONAL VET PROJECTS

The transformation of VET ecosystems requires time, resources and changes in behaviour and governance. Promoting local ownership can increase commitment and adoption while providing the evidence needed for system actors to replicate the success in other regions and sectors. Swisscontact's project experience demonstrates that for VET systems to respond to social and economic needs within a dynamic environment, ecosystem actors need to operate within an enabling local environment and have the relevant capabilities and interests to redefine their roles and responsibilities and drive change. Creating such a situation requires the adoption of systems thinking and practice approaches.

   A conventional VET project predefines key partners, focuses narrowly within the VET system on the delivery of training to target groups and places most of its attention and decision-making power on government actors and public providers, with a smaller emphasis on private providers and promoting industry engagement. This tends to isolate the VET system and create fragmentation by downplaying the wider ecosystem. In contrast, ISD considers all the actors within an ecosystem by analysing the overlapping subsystems, such as specific industry segments in different local areas. The approach acknowledges the roles and functions of both public and private players, including the crucial role that government must play in overseeing the VET sector (such as regulating, assuring quality, etc.) and in acting as an accelerator of change. This requires an emergent and adaptive approach towards building relationships and

negotiating partnerships between government, VET providers, employers, learners and other actors. It becomes in their best interests to work together while promoting shared value between stakeholders. Typical actions seek to sharpen the incentives or build the capability of system actors to address bottlenecks by engaging in new ways of working.

Thus, there is a tension between a conventional VET project and an ISD approach. For example, ISD requires different skill sets and competencies from project staff to encourage interconnectedness and facilitate multistakeholder engagement over a longer time horizon by using a flexible and adaptive approach. This includes a shift in appetite for risk and experimentation from the donor, the implementing partner, project staff and ecosystem actors themselves – who may be comfortable with the predictability of a more traditional and linear approach. Nonetheless, the challenge is not insurmountable, and the MSD approach, which has been applied over many years in differing contexts, offers experiences and tools for its application in VET ecosystems.

ISD-SD: DRAWING LESSONS FROM MARKET SYSTEMS DEVELOPMENT

The MSD approach has evolved over 20 years (Committee of Donor Agencies for Small Enterprise Development, 2001; Ledgerwood, 2021; The Springfield Centre, 2008, 2015). MSD grew out of economic development and private sector development programming and is built on the assumption that system dysfunctions have a disproportionate impact on poor and marginalised groups but that the system itself can ultimately overcome these via effective analysis, flexibly co-created interventions and facilitative actions. The approach determines a highly flexible set of 'system boundaries' that allow analysts to zoom out, reduce the complexity and zoom in on the key nodes of influence. MSD analysis encourages movement between the core system, interconnected supporting systems and enabling environment to delve into the root causes of system dysfunctions and underperformance that impact the poor. MSD promotes facilitative actions that catalyse behaviour change. In practice, these usually involve a combination of risk sharing, technical assistance and negotiation tactics to convince market actors to invest in new business models based on calculated expectations of benefits. The new business models are then intentionally spread across an entire market system while being replicated in neighbouring systems. This foregrounds the sustaina-

bility of behaviour changes and builds on a goal of reaching scale through a diffusion of innovations that improve outcomes for target groups.

An MSD system perspective offers lessons to ISD-SD when it comes to overcoming fragmentation and developing a common understanding and ownership for skills development through an ecosystem perspective, which encourages stakeholder engagement processes with a focus on sustainability and scale. By taking a less direct approach to project delivery, innovations can be anchored more firmly in local systems, and change can accelerate beyond project boundaries and partners. MSD projects allow time for trust and confidence to emerge, relationships to strengthen and cocreation processes to unfold. MSD is enabled by supportive donors who encourage projects with an extended inception phase, which can help deepen their systemic analysis based on the view that durable change must address root causes and that it takes several years to build relationships and ensure behaviour change.

MSD offers an adaptive management practice and a set of refined tools for assisting project staff in facilitating systemic change to achieve impact at scale. It also provides proven strategies and tactics for developing and negotiating strategic partnerships with local actors that ensure the ownership of change processes within a system while assessing power dynamics and governance issues at different levels. These elements can bring practical benefits to ISD for skill development projects.

Learning from these experiences opens up an opportunity to see how VET ecosystems can be approached. However, the framing of a 'core market' – with supply and demand as formulated in MSD – does not neatly translate into the holistic view of lifelong learning for all in SD (which includes VET). ISD-SD requires a more normative perspective because VET systems seek to fulfil multiple goals: rights-based access and inclusion; skills shortages and skills gaps identified by the industry to address productivity and competitiveness; and broad development goals, such as inclusive and equitable growth. The focus is on VET as a system and how it needs to be strengthened, with its own actors involved and linked to a variety of ecosystems (education, economy, labour markets, etc.), which have different goals and interests that need to be met and negotiated.

Therefore, it is important to critically examine the analytical tools and instruments drawn from the MSD experience to make relevant adjustments when applying them in a SD context. This recontextualisation of ideas from one sector to another requires sustained engagement and experimentation from practitioners in both spaces. Over the course of

several years, a cross-functional team within Swisscontact has derived four core ISD principles that underpin project design and implementation in different contexts that work together to deliver inclusion (see also Weyer & Kehl (2015) for findings from a Swiss Agency for Development Cooperation internal working group).

ADAPTING ISD PRINCIPLES TO SKILLS DEVELOPMENT CONTEXTS

This section discusses the key principles of Swisscontact's ISD approach and how these principles are interpreted and applied in the context of ISD for skills development (ISD-SD). A case study of a SD project is used to show how ISD principles are integrated into the design and implementation of the Skills Development Programme II in Cambodia.

The four ISD principles have been derived from an internal synthesis conducted by Swisscontact based on MSD best practices, mostly in agriculture projects. The aim was to provide an approach that encourages systemic change in multiple sectors of development, including where system boundaries are unclear, where there is high interconnectedness of ecosystems and subsystems, and where social outcomes are just as important as economic outcomes. In this section, we pay particular attention to the differences in how the core principles are interpreted and can be applied in an ISD-SD project.

Systems Thinking

To ensure enduring systemic change, the underlying causes need to be addressed. This requires understanding systems in terms of the roles and relationships between actors and institutions and their interconnection to other subsystems within a broad ecosystem. It also involves taking a big picture view to determine the entry points for unlocking the system to deliver equity and inclusion and to ensure that benefits reach marginalised actors. This can be done by asking questions such as the following: Why is this behaviour occurring in the ecosystem? Why do SD services not reach marginalised groups? Why is the system itself not resolving the issue?

A primary adaptation for ISD-SD when it comes to applying systems thinking is to draw system boundaries that include a wide enough set of actors.

BOX 8.1   DEFINITIONS OF 'SYSTEM' IN TWO THEORETICAL APPROACHES

MSD	ISD-SD
Primary systems of interest are market systems, which are defined in terms of supply and demand for a product or service. Systems analysis locates the target group and determines dysfunctions and market failures that impede its participation or performance in markets. Typically, dysfunctions are found in the interconnected systems of a core market – this could be an enabling environment issue, a social norm, key infrastructure, unequal market power or lack of information or a particular skill.	VET ecosystems comprise multiple interconnected subsystems, including education, initial VET, continuing VET and labour market institutions like employment services with their own logics and interests. VET systems must fulfil multiple goals: at the lower level, more public VET focuses on access and inclusion for all, including disadvantaged groups, while higher-level VET operates with a market logic to address specific skills gaps so that enterprises remain competitive. Target groups can be difficult to locate in VET ecosystems because they 'move' through different subsystems and optimise their SD and use (e.g., education, labour market insertion services, employment or self-employment).

In practice, the challenge when applying ISD-SD is to coherently integrate these distinct 'subsystems' and facilitate their effective interaction to achieve optimal results, typically productive employment and self-employment. This includes how they are framed, funded and regulated by government agencies (often there is poor coordination between the Ministry of Education and the Ministry of Labour or their equivalents) and how well they are coordinated jointly with the private sector. One practical approach to visualising subsystems is to develop multiple parallel system maps to emphasise actors, relationships and functions in different ways.

Ecosystem Facilitation

To facilitate a process and build the capability of the system to function more inclusively and effectively, collaboration should be fostered. This is achieved through catalytic actions that support ecosystem actors to understand each other, change behaviour and foster local and national cocreation. Facilitation is a stance and identity for time-bound development projects and a reminder that change needs to be owned by local

actors. It involves asking questions such as the following: How can we build capabilities and enable local system actors to drive change based on their own interests and consensus with other actors? How can we reinforce promising new patterns of behaviour at the local and national levels?

The main adaptation for ISD-SD lies in understanding the broader range of non-economic incentives and potentially conflicting interests of system actors while factoring this into the complexity of partnerships and relationships being facilitated.

BOX 8.2   TWO APPROACHES TO FACILITATION WITHIN A SYSTEM

MSD	ISD-SD
Market facilitation is the set of activities used to stimulate new behaviours and relationships between market actors. Anchored in strong business analysis, market facilitators negotiate new types of partnerships by coaching market actors to try new business models. The facilitator brokers targeted relationships and specific partnerships to deliver a new way of working – but often not in a purposefully participatory manner.	ISD-SD creates spaces and platforms at national and subnational levels for a wide breadth of actors to come together to understand each other's roles and relationships and to analyse and co-create a transparent process for engagement and solutions for change. VET systems are highly regulated, with the result being that change needs to be negotiated and is slow. Facilitation occurs at multiple levels: (1) work with individual VET providers to change their strategies and educational offerings to ensure access and inclusive services to more potential learners and people in need for upskilling and reskilling; (2) facilitate two-way relationships between employers, employment services, government agencies and VET providers to ensure relevant skills and labour market insertion; and (3) facilitate multistakeholder platforms that build trust, share information and support joint action by key players as a way to design and deliver relevant services.

Working at multiple levels requires clear strategic thinking from managers and adaptive management skills from project staff. This can be particularly challenging in practice because SD projects are often con-

strained by a prescriptive project design that limits the space for adaptation and selecting partners on an ongoing basis.

Sustainability

Sustainability is defined as enduring behaviour change within and between actors that last beyond a project's lifetime. This requires a clear and realistic vision for the future functioning of the ecosystem, including roles, relationships and interconnections. Sustainability builds on the capacity and incentives of ecosystem actors (both public and private) to adopt innovations and own and drive change, and it involves asking questions such as the following: How can we ensure that system actors continue the new ways of working after the project ends? Who will pay for this function in the system in the short and long term?

Ecosystem facilitation is the means to achieve the end goal of sustainability, here based on fostering ownership by local actors. Again, the adaptation builds on truly understanding the incentives of the full range of skill development actors.

BOX 8.3   HOW THE TWO THEORIES APPROACH SUSTAINABILITY

MSD	ISD-SD
Innovations are embedded based on actor incentives to change – this is often via business models owned by private sector actors but not exclusively so. New business models become entrenched in industry when multiple competing actors adopt them – pushing each other to keep adjusting and innovating based on a competitive drive. MSD tends to promote market-based solutions and a stronger inclusion of private actors and market mechanisms.	ISD-SD projects need to grapple with non-economic incentives for diverse actors to jointly develop ownership and commitment for change. Public financing is crucial for sustainability of inclusion, participation and access, and access to this financing depends on a complex set of rules, regulations and requirements (e.g., accreditation standards). Public VET providers respond to a diverse set of incentives – achieving political and social outcomes – which are seen as legitimate educational providers for quality VET, meeting the needs of students and parents and justifying public spending. At the same time, private VET providers and employers investing in training may still be driven by commercial incentives while operating within the rules and fulfilling the requirements to be accredited.

In practice, a focus on sustainability requires the tools and frameworks that make explicit what behaviour needs to change, the incentives for ecosystem actors to make change and the financial sustainability of those changes. A widespread MSD tool for this analysis is the 'who does, who pays' matrix (BEAM Exchange, 2021) which compares the current and future set of roles of different actors in the system. This has been modified and used successfully in ISD-SD projects.

Scale

Scale is defined as developing strategies that expand successful innovations to reach large numbers of people in target groups through involved partners and actors. Scale can be achieved through different mechanisms, such as (a) policy changes that mandate the adoption of the change or (b) stimulated replication where change is adopted by similar organisations and actors in an ecosystem. Scale involves asking questions such as the following: How can successful pilots scale up beyond project support? What are the pathways or mechanisms for this change to become more widespread?

In ISD-SD, scale is interpreted as reaching many public and private training service providers to impact large numbers of learners or workers. Adaptations arise from considering different mechanisms for reaching scale and working with significantly more regulations and quality assurance.

BOX 8.4   HOW THE TWO APPROACHES ADDRESS THE ISSUE OF SCALE

MSD	ISD-SD
Creating change in one supporting market (e.g., the market for input supplies for agriculture) has a ripple effect on a wide range of interconnected markets (e.g., for different crops or value chains). The main innovation being scaled is a business model, and the mechanism for achieving scale is 'crowding in' – the copying and adaptation of new business models by multiple actors within one market system and expansion into other market systems. Thus, competition between firms is a key mechanism for reaching scale.	In ISD-SD, the innovation being scaled can take a number of different forms. Innovations could be short- to long-term school-based programme curriculum and pedagogy; nonformal or formal in-company learning structures; occupational profiles, standards, assessment and certification; national, sector or local governance structures; or linkages between education and labour markets. Similar to MSD, sometimes these models are adapted and scaled through replication (e.g., VET schools copying each other's curricula or companies adapting an internship or apprenticeship model of their competitors). A key difference is the central role of the local and national public sectors as key accelerators for achieving scale. For example, in public VET provision, projects could partner with leading VET schools or networks, as well as government ministries, to support changes to policy that enable (or force) other public or private providers to adopt new workplace-based practices or curricula, teacher training and assessment.

A key difference is the central role of the local and national public sectors as key accelerators for achieving scale. For example, in public VET provision, projects could partner with leading VET schools or networks, as well as government ministries, to support changes to policy that enable (or force) other public or private providers to adopt new workplace-based practices or curricula, teacher training and assessment.

In practice, ISD-SD projects need to be clear on what innovations they want to scale, and what mechanisms offer the best chance to reach scale. The project design needs to be aligned with the dynamics of the VET ecosystem – strategies need to be adapted to the local context, and there needs to be close partnerships and buy-in for policy changes to be realistic. Having a clear picture of the range of mechanisms and clear

understanding of the roles and responsibilities of each actor in design, implementation and financing gives project teams more options for achieving scale. This is a time- and resource-consuming process that requires a long-term programming perspective with clear vision and ownership for change.

CASE STUDY OF ISD IN THE SKILLS DEVELOPMENT PROJECT IN CAMBODIA

The four ISD principles overlap and are interdependent; they also need to be applied in a coherent and integrated fashion. To demonstrate this, we discuss the example of the Cambodia Skills for Development (SDP) II project, an ongoing skills reform project of the Swiss Agency for Development and Cooperation (SDC) that is implemented by Swisscontact in consortium with the Institute for Vocational Training, Labour Market and Social Policy GmbH (INBAS). The project's high-level goal is to create decent wages and self-employment that can lead to increased income for disadvantaged and marginalised women and men; it proposes to do this through improving the effectiveness of the national TVET[1] system and selected provincial TVET ecosystems.

Phase 1 of the Skills Development Programme was designed as a conventional project: reforming curricula and partnering directly with provincial TVET schools and private training providers to achieve significant training and employment targets. A lack of attention to overall governance structures of the ministry at the macro, meso and micro levels limited the sustainability of partnerships. Based on the midterm review, the project shifted to a collaborative approach to better integrate interventions into partner structures. This supported a shift towards an ecosystem facilitation approach in Phase 2 to build more ownership among TVET system actors to achieve deeper systemic change. This also includes institutionalising change through the commitment, ownership and guidance of key stakeholders within steering structures. The project's technical team made ISD a guiding management approach by continuously training the team and reinforcing ISD tools and instruments. In this way, interventions could be integrated into the management system,

---

[1]   In Cambodia, the term TVET (technical and vocational education and training) is used instead of the term VET, so we use TVET throughout the case example for accuracy.

which constantly evolves and is tailored to the strategies and operational processes of stakeholders in the TVET system.

Systemic Challenges

The TVET system in Cambodia suffers from typical challenges that many countries face: it is a small, uncoordinated and supply-driven system that is weakly connected with industry. Limited collaboration between system actors is most pronounced between the government and private sector employers, but also between government ministries. Significant donor funding and active international NGOs play a role in delivering training directly, which can undermine the role and capacities of public TVET providers, such as provincial training centres (PTCs). This contributes to a fragmented sector with weak capacity and low quality and relevance of training, both in public TVET schools and in companies. The overall system lacks financial resources, practical procedures, legislation, staff and time to successfully run legal entities, such as the Sector Skills Council (SSCs).

The project's vision of change (Figure 8.1) addresses these challenges through four main intervention areas. Each intervention area (e.g., intervention area 1 (IA1), TVET school development of PTCs at provincial level) targets a change in behaviour by those actors delivering services within the system ('System trigger'), which is intended to stimulate a response ('System uptake') that will have an ultimate impact on the project's target groups. At the surface level, this looks like a straightforward logic model. Systems thinking becomes more visible in the adaptation of ISD tools to analyse the relationships between actors within the TVET system. This is represented in two system maps.

The skills actor map (Figure 8.2) shows the main actors – government, industry and TVET training providers – and the nature of relationships between them. The depiction of the government shows the different agencies involved (e.g., National Employment Agency, Ministry of Labour and Vocational Training and Director General of TVET) at both the national and provincial levels and the lack of coordination between them. Each potential relationship between actors is characterised; for example, the poor link between industry and SSCs. This paints a picture of the dynamics of the current system – and many of its problems, including fragmentation, a lack of coordination and disconnects between labour market needs and target group skills. The skills actor map is crucial for
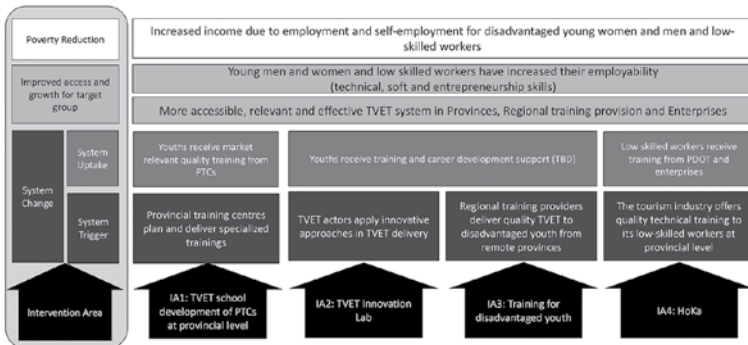
Figure 8.1   SDP Cambodia's vision of change

building a shared understanding with the system actors themselves, as well as building credibility with sectoral experts in vocational education.
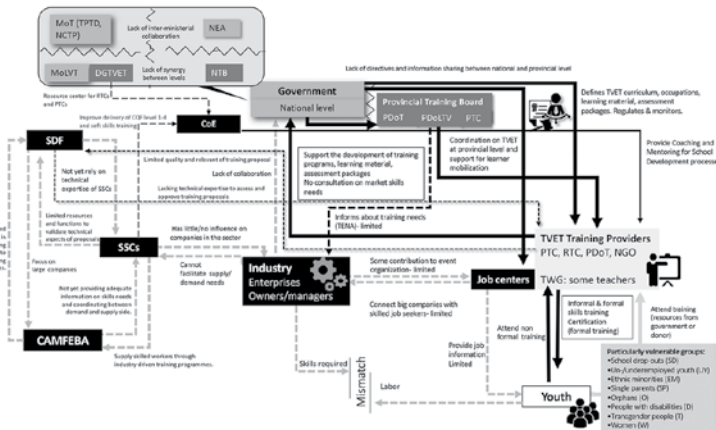
Figure 8.2   System map focused on key actors

An additional lens is provided by the system functions map (Figure 8.3), which focuses on the core service chain between labour supply and

demand and the underlying support and regulatory functions that are the root causes of problems in the TVET system. These are numbered (e.g., SF1 or RF11) to keep the project team focused on the key issues and ensure each intervention is addressing a root cause.
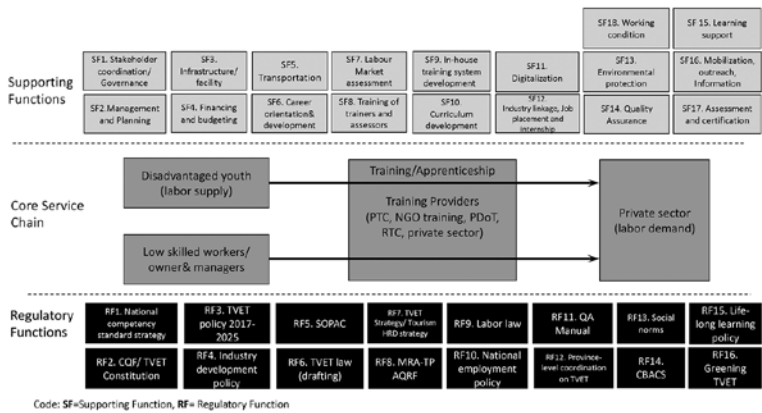
Source: Authors.

Figure 8.3   System map focused on key functions

Together, the two system maps paint a picture of the structure and complexity of the overall TVET system in Cambodia. The maps were initially developed based on primary data collection and the experience of an earlier phase of the project. In shifting to the second phase (the current phase), the maps were refined further through one-on-one meetings with many of the key stakeholders to understand how well they represented the current realities of the system. These meetings helped to go from a broad characterisation of relationships between key actors (e.g., strong/ weak) to a more nuanced articulation of the two-way relationship. For example, SSCs have limited resources to review proposals from the Skills Development Fund (SDF) and, in turn, the SDF does not yet rely on the technical expertise of SSCs.

   In addition to the overall system maps, the project has developed analyses for its four intervention areas. Figure 8.4 shows the current model for intervention area 1 (IA1) – public TVET provision through PTCs – with crucial service and regulatory weaknesses labelled at the intersection of key actors. There are mismatches between what PTCs

offer, what disadvantaged youth want and what industries need. This is underpinned by other disconnects within the government structure – from national policies to provincial job centre priorities.
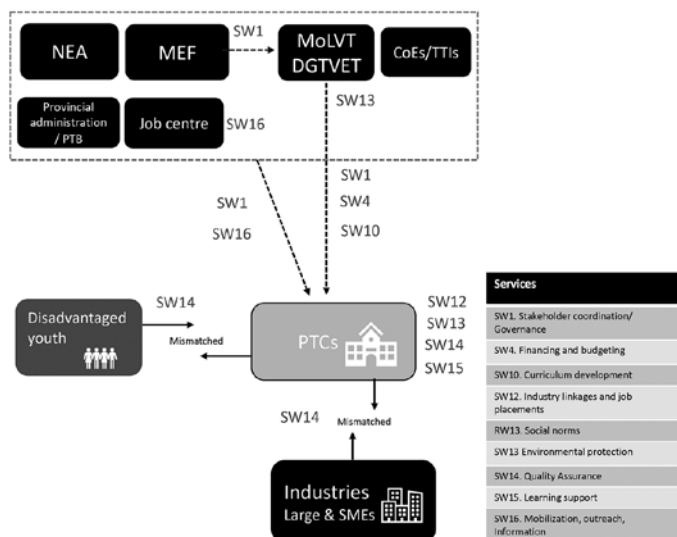
Figure 8.4   Current model for intervention area 1

A shared vision of a desired future model for the PTC subsystem in intervention area 1 is shown in Figure 8.5. The new model focuses on the directorate general's vision to reshape PTCs to become institutes that function in a dynamic relationship with key local industries, as supported by an integrated and coherent set of supports from the Ministry of Labour and Vocational Training (MLVT). The relationships between government actors are clarified, as are roles and responsibilities at the provincial and national levels. The model depicts a different set of relationships between the key actors and links to the crucial services from the functions map – articulating the desired direction of change. From a sustainability perspective, this firmly reflects the thinking that system actors must be engaged in the process because they need to change behaviours and relationships. The Swisscontact project acts as a systems facilitator to work with those actors to make changes.
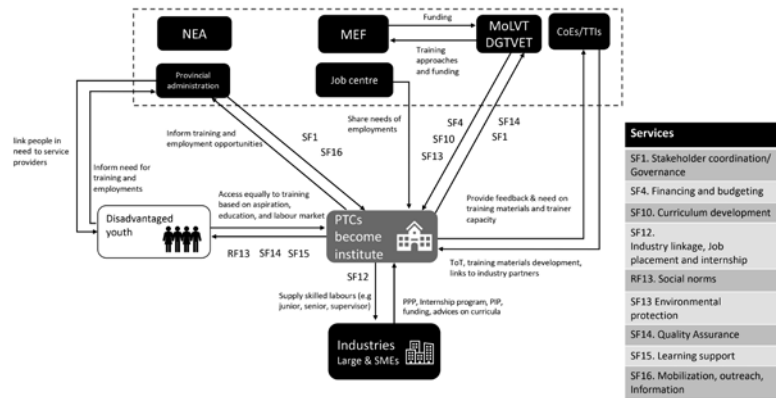
Figure 8.5   Vision of a new model for intervention area 1

Mirroring the current system with a vision of a desired future system, it is possible to isolate facilitative actions that can change the behaviour of actors towards the vision of a new model and address the weaknesses in core services, support and regulatory functions. Key services can be further broken down into (a) the underlying causes of the weaknesses, (b) activities to instigate change and (c) the new responsibility of the system actor. An example (Figure 8.6) is shown for the service function 'quality and relevant TVET trainings', which is at the core of PTC offerings.

With a clear systems analysis and direction on what needs to change, the project staff can work directly with system actors to develop partnership agreements and road maps to drive change based on local ownership. The SDP's partnership agreements with PTCs are tripartite agreements that ensure (a) management of the PTC retains ownership of changes, including commitment of internal resources; (b) project support is temporary with clear exit plans; and (c) reinforcement and support of these from the government director general of TVET. One effective tactic has been to use the internal quality assurance process of MLVT to leverage the appropriate government mechanisms that reward changes in the service offerings of PTCs.
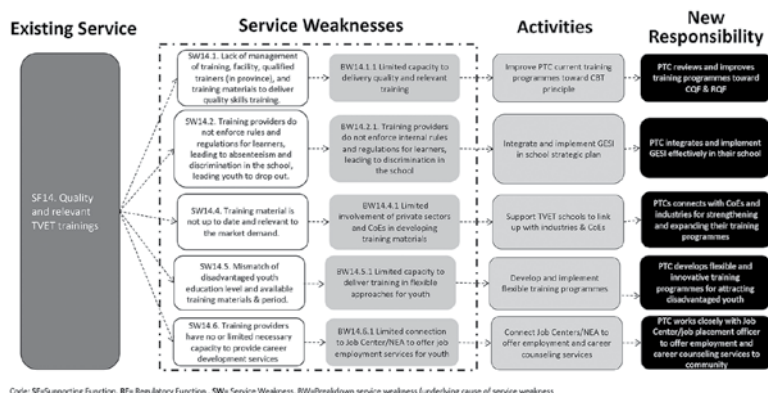
Figure 8.6   *Change table for intervention area 1, with a focus on SF14, service weakness for quality and relevant TVET trainings*

KEY LESSONS FROM CASE STUDY

This project is being implemented at time of writing, but key lessons have already begun to emerge.

First, a significant sustained effort from project leadership is required to build ownership and capacity for taking an ISD approach to skill development. This extends far beyond training project staff: ongoing mentorship from an in-house team of experienced ISD leaders is critical to help project teams internalise the logic of a systemic change strategy and adapt the tools to work for their intervention area. Having the same mentors to lead training, walk through intervention planning with team members and adjust tools where needed is a major asset and is more effective than bringing in external consultants. Facilitation can be uncomfortable at first for staff used to working on projects where there is direct delivery of training services. However, with support and encouragement they have come to see the benefits in terms of local ownership and, ultimately, increased sustainability. In terms of the composition of the team, a mix of skills and experiences supported by a committed team leader is key to reinforcing the importance of ISD in every meeting,

with committed mentors available to problem-solve field-level problems regarding applying ISD.

Second, access to timely information on the status of different changes is crucial to decision making and course correction. To support this, the project has an adaptive monitoring and results measurement (MRM) system that assesses and articulates the intended sequence of changes for different actors, with qualitative and quantitative indicators to track the extent of buy-in and adoption. The team then uses periodic 'deep-dive' meetings to present the latest data to key stakeholders and reflect on whether interventions need to change based on the findings. This includes revisiting the results chains and change table to ensure that the underlying weaknesses are being addressed. Having clear structures and processes on which to base any changes in strategy and tactic is important for transparency and for gaining support from the donor to make those changes.

The third lesson is that having a steering structure comprised of national and provincial actors helps to gauge progress and keeps the relevant authority (DGTVET) abreast of any major changes; this ensures these actors have the support required. The steering structure was initiated by the project as a temporary measure, but one that can become institutionalised beyond the life of the project to continue supporting the key interventions and to keep the PTCs connected to national-level policymakers. Significant effort has been invested in developing and supporting steering structures with participation from the key stakeholder groups: ministries, VET schools, industries and disadvantaged learners themselves. These steering structures, combined with the MRM system, help drive adaptive management: reflection on progress that informs decision making about changes.

IMPLICATIONS FOR DEVELOPMENT COOPERATION AND POLICY

Although this chapter has focused on the VET system, it has important implications for the uptake of systems approaches to VET in development cooperation and policy more broadly.

A clear definition of the system(s) is a crucial foundational point. The SDP example shows the benefit of representing systems in multiple ways: in terms of actors and relationships, but also in terms of functions. By analysing the position of excluded groups and stakeholders within a system map and their relationship to more powerful actors, it is easier to

identify the underlying dysfunctions and weaknesses that hold them back from meaningful participation, capacities and ownership.

To operationalise a systems approach, it is also crucial to have adaptive management structures and processes in place that enable an ongoing evidence-based and transparent dialogue for the revision of assumptions and to adjust to changing dynamics. One component of such a system is a clear articulation of what should change in the VET or education system – actors and their roles, relationships, collaboration and functions. From here, it is possible to gather information to track the extent to which such changes are happening and to allow course correction and reflection. For change to be owned locally, it is important to create well-defined and transparent mechanisms that allow various system actors to come together to critically reflect and jointly make decisions on progress based on data regularly collected.

There is rarely a development initiative that is given the full freedom to adapt and learn without any semblance of predefined log frames, budgets, timelines or targets. The SDP Cambodia example shows an evolution from a first-phase project design that resulted in a focus on the number of people receiving training to a more flexible second-phase design emphasising sustainability through ownership and reinforcement by local system actors themselves. This is a result of an effective and self-critical midterm review process and extended dialogue between the donor and implementing organisation. Other development cooperation initiatives seeking to stimulate systemic change in education systems would benefit from proactive dialogue. This would include a discussion of the implications of setting ambitious and short-term training targets linked to overly detailed budgets, which can erode the longer-term ownership of change processes by local actors.

Finally, the example highlights the importance of assembling broad project teams with the right mix of sectoral expertise in education and process or methodological expertise in systemic change. An ongoing commitment to mentorship and culture of learning are also key to bringing a group of people along in facilitating systemic change.

CONCLUSION

ISD-SD offers a promising approach for creating systemic change in VET ecosystems. By empowering actors from across the wider ecosystem to take on new roles and responsibilities and form new relationships and practices, ISD's four principles of systems thinking, sustainability,

facilitation and scale provide a framework to overcome continuous failures in VET system reform. Several tools and processes for ISD have been drawn from related MSD processes. However, they require adaptation to the VET context because of differences in normative goals, strong regulations and contextual features of VET systems, along with different requirements and norms of the organisations that fund and implement VET and SD projects.

Internalising the ISD-SD approach requires a suite of changes in how VET policymakers, donors and project implementers operate. From a project design and implementation perspective, attention needs to shift away from a narrow focus on addressing reforms of VET systems alone and the currently dominant focus on government-driven and controlled VET to instead understanding how the system creates barriers to the participation of actors, including disadvantaged groups. Projects should be framed around these critical system dysfunctions and offer enough space during the inception phases for iterative systems analysis and a deep and wide stakeholder engagement process that uncovers the incentives and capabilities of local actors to drive different types of change. Changes in culture and behaviour require time, money and patience but will yield major benefits in terms of sustainability and the prospects for more private sector engagement and ownership, as well as scale and replication across systems.

We point to the success of the MSD approach in coalescing a critical mass of NGOs, contractors, donors and consultants to agree on the principles, tools and practices over a period of two decades. Given the importance of changing education and VET systems to ensure the successful futures of so many people across the world, it is worth the effort to build a similar coalition specific to education and skills development.

REFERENCES

Allais, S. (2020). Skills for industrialisation in sub-Saharan African countries: Why is systemic reform of technical and vocational systems so persistently unsuccessful? *Journal of Vocational Education & Training*, 74(3), 1–19.

BEAM Exchange. (2021, 15 September). Who does, who pays? https://beamexchange.org/guidance/vision/who-does/

Bennell, P. (1996). Rates of return to education: Does the conventional pattern prevail in sub-Saharan Africa? *World Development*, 24(1), 183–199. https://doi.org/10.1016/0305-750X(95)00115-S

Blattman, C. & Ralston, L. (2015). Generating employment in poor and fragile states: Evidence from labor market and entrepreneurship programs. SSRN 2622220.

Buchanan, J., Anderson, P. & Power, G. (2017). *Skill ecosystems*. Oxford University Press.

CEMETS. (2021). CEMETS Education Systems Reform Lab. https://cemets.ethz .ch/the-center.html

Committee of Donor Agencies for Small Enterprise Development. (2001). Business development services for small enterprises: Guiding principles for donor intervention. Donor Committee on Enterprise Development (DCED). https://www.enterprise-development.org/wp-content/uploads/BDS-Guiding -Principles-2001-English.pdf

European Training Foundation. (2021). Torino process. https://www .torinoprocess.eu/

Finegold, D. (1999). Creating self-sustaining, high-skill ecosystems. *Oxford Review of Economic Policy*, 15(1), 60–81.

Ledgerwood, J. (Ed.). (2021). *Making market systems work for the poor: Experience inspired by Alan Gibson*. Practical Action Publishing.

McGrath, S., Powell, L., Alla-Mensah, J., Hilal, R. & Suart, R. (2020a). New VET theories for new times: The critical capabilities approach to vocational education and training and its potential for theorising a transformed and transformational VET. *Journal of Vocational Education & Training*, 1–22.

McGrath, S., Ramsarup, P., Zeelen, J., Wedekind, V., Allais, S., Lotz-Sisitka, H., Monk, D., Openjuru, G. & Russon, J.-A. (2020b). Vocational education and training for African development: A literature review. *Journal of Vocational Education & Training*, 72(4), 465–487.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? A critical review of recent evidence. *The World Bank Research Observer*, 32(2), 127–154.

Okumu, I. M. & Bbaale, E. (2019). Technical and vocational education and training in Uganda: A critical analysis. *Development Policy Review*, 37(6), 735–749.

Openjuru, G. L. (2010). Government education policies and the problem of early school leaving: The case of Uganda. In *The burden of educational exclusion* (pp. 15–33). Brill Sense.

Powell, L. & McGrath, S. (2019). *Skills for human development: Transforming vocational education and training*. Routledge.

Psacharopoulos, G. (1981). Returns to education: An updated international comparison. *Comparative Education*, 17(3), 321–341. https://doi.org/10.1080/ 0305006810170308

Psacharopoulos, G. (1985). Returns to education: A further international update and implications. *Journal of Human Resources*, 20(4), 583–604.

Rekiso, Z. S. (2019). Education and economic development in Ethiopia, 1991–2017. In *The Oxford handbook of the Ethiopian economy* (pp. 428–446). Oxford University Press.

Swisscontact. (2021). Skills development in a changing world. Swisscontact.

https://www.swisscontact.org/_Resources/Persistent/7/d/a/1/7da11eb4258
4e9859f1b8a0a4e18d4ac73850126/Swisscontact%20Skills%20Development
%202021.pdf

The Springfield Centre. (2008). *The operational guide for the making markets work for the poor (M4P) approach*. The Springfield Centre.

The Springfield Centre. (2015). *The operational guide for the making markets work for the poor (M4P) approach* (2nd edn). The Springfield Centre. www.springfieldcentre.com/wp-content/uploads/2015/11/2015-09-M4P-Op-Guide-Sept2015.pdf

Weyer, F. & Kehl, F. (2015). SDC working group on vocational skills development and the market systems development approach. Synthesis. Swiss Agency for Development Cooperation.

9. Systems, complexity and realist evaluation: reflections from a large-scale education policy evaluation in Colombia

**Juan David Parra and
D. Brent Edwards Jr**

INTRODUCTION

One central contention behind the writing of this book is that although systems thinking (ST) has permeated disciplines like medicine, agriculture and urban studies, its influence in education is still mild. From a policy analysis and evaluation viewpoint, a potential (structural) explanation for this situation might find its roots in what has been labelled the political economy of knowledge production and mobilisation in the global education policy field (Edwards et al., 2020; Broad, 2006). That is to say, the significant investments in recent decades by organisations such as the World Bank in certain types of studies that fit into their organisational preferences (e.g., regression analysis, randomised controlled trials) provide few incentives for the emergence of alternative methodological approaches among education policy researchers (Edwards, 2018). Whatever the more specific explanation might be for better understanding why ST has not spread more in the analysis of educational interventions, the reality is, as Spivack (2021) confirms, that discussing ST with policymakers and practitioners often induces 'eye-rolls and groans'; she argues that this is the case because the topic is often perceived as 'too theoretical' and, hence, 'far removed from the practical, pressing concerns facing country-level decision-makers and implementers' (p. 2).

This chapter consists of a methodological reflection that contributes to the debate on how to incorporate ST concepts and principles into

the assessment of education policies and interventions, hence inviting different audiences to see the merits of this approach in informing education policy debates. The ideas we discuss emerged in the context of the evaluation of the implementation process of a large-scale educational intervention in Colombia between late 2018 and mid-2019.[1] The Colombian Full-Day Schooling programme (or JU, for its acronym in Spanish[2]) has been described by researchers from the Organisation for Economic Co-operation and Development (OECD) as 'an opportunity to test an inter-institutional coordination model (…) to plan and implement various multiyear school improvement strategies' (Radinger et al., 2018, p. 118). Therefore, our assessment of JU entailed navigating between the processes and social interactions happening simultaneously at different levels (e.g., the national, subnational and local) of the Colombian education system. Hence, the experience we report here is a testimony to the feasibility and advantages of thinking systematically about educational interventions in the context of non-academic projects and under the budget and time constraints of decision-oriented policy research.

From an epistemological perspective, Caffrey and Munro (2017) have argued that '[t]he application of systems approaches to evaluation can be categorised as a form of theory-based evaluation in the same family as (…) Realist Evaluation [(RE)]' (p. 465). Other authors agree with this appreciation of the significant similarities between the guiding principles of (at least some approaches to) ST[3] and the realist approach to the assessment of social interventions to the point of presenting them as 'natural bedfellows' (Westhorp, 2012, p. 407). However, to the best of our knowledge and although highly popular in disciplines such as medicine or criminology, RE has not decidedly penetrated the literature and practice of policy assessment in the field of education (Tikly, 2015). Hopefully, the ideas conveyed in this chapter – in which we discuss some technical

---

[1]   Note that only the first author of this chapter participated in the evaluation, but to improve the flow of the narrative, the language throughout uses the first person plural (we, us, our).

[2]   The original name of the programme in Spanish is *Jornada Unica*.

[3]   The history of ST shows the existence of different epistemological standpoints to the study of natural and social systems (Turner & Baker, 2019; Mingers, 2014; Midgley, 2006). However, the epistemological grounds of RE, whose exponents associate it with the tenets of scientific realism (Jagosh, 2020; Pawson, 2013), evoke the foundations of complexity theory (Mingers, 2014; Cartwright, 2013; Byrne, 2005).

details of our evaluation design of JU and reflect on our experience in applying them to inform our fieldwork and data analysis decisions – will contribute to changing this situation.

We have split the chapter into three sections. First, we expand our discussion on the parallels between ST and RE, emphasising the tools of the latter, which, in our view, helps make abstract systems' principles and concepts usable in a practical evaluation setting. Afterwards, we document our policy assessment experience, accentuating our methodological decisions made during different stages of the evaluation process. These stages involved the following: defining the boundaries (or components) of the system we were exploring to focus the data collection process, sampling subregions of the country and schools to visit to explore different system's dynamics, establishing data analysis protocols and conveying findings to policymakers and implementers.

FLESHING OUT SYSTEMS THINKING PRINCIPLES: THE REALIST EVALUATION APPROACH

One main argument for taking ST seriously in evaluation studies is that 'system reform [in education] cannot be effective unless "how and why" the system functions is taken into account' (Gillies, 2010, p. 34). Such a claim might still sound too abstract for many experts and practitioners in the field of education policy, who might even question if the role of policy or programme assessment is to address system- or structural-level reform. However, as Midgley (2006) contends, '[in] the context of evaluation, it is widely recognised (…) that setting narrowly defined goals for a service or organisation, and measuring the achievement of these alone, may result in the evaluator missing positive benefits that lie outside the scope of the evaluation' (p. 11). From this perspective, evaluations will hardly provide sound guidance for future policy/programme implementation decisions if they omit the analysis of elements that might operate beyond the confines of an intervention's planned processes and activities but are nevertheless central in explaining its different outcomes (Moore et al., 2019).

Broadly defined, a system is a 'set of entities (e.g., people, organisations, resources) and their interconnections' (McGill et al., 2021, p. 2). In the realm of evaluation, we can argue 'that programmes [and policies] are complex interventions introduced into complex social systems' (Pawson, 2013, p. 33). This means that their implementation

happens in contexts with pre-existing features, such as specific material (e.g., distribution of financial power) or cultural (e.g., dominant ideas) backgrounds, that condition, but do not determine, people's reactions and interactions in response to the new intervention (Jagosh, 2020; Pawson, 2013). Consequently, studying the drivers of different outcomes of educational interventions entails addressing how their recipients – that is, the inhabitants of the social system in question – reflect and act upon the resources (e.g., subsidies, training in teaching skills) delivered to them (Dalkin et al., 2015). Such enquiry logic differs quite notably from the more traditional input–output analyses of mainstream impact evaluation models that aim at corroborating statically significant connections (or causal ascriptions) between variables (Joyce & Cartwright, 2020).

However, applying the principles or ST-based considerations mentioned above in evaluation settings entails certain methodological challenges for evaluators. As Byrne (2005) puts it, '[the] big question is how can we interrogate [social systems] to understand how things [or specific programme outcomes] have come to be as they are and how they might be made different' (p. 101). Social interventions seek to induce some change (e.g., improve student's school performance). Still, those changes – because of uncertainty in how people will understand the intervention and react to it within the specific contextual features of the social system they inhabit – 'can never be fully anticipated and [are] not entirely predictable' (Pawson, 2013, p. 6). In the ST literature, scholars use different terminologies (e.g., bifurcation, emergence and feedback loops) to conceptualise such complexity in the study of system dynamics. However, practitioners often critique these frameworks for being too 'wordy' and 'too abstracted' (Caffrey & Munro, 2017, p. 464).

In line with our suggestion in the introduction, the language of RE and its enquiry rationale provide practical and methodologically sound tools – compared with other (abstract) academic proposals – that can help bridge the theory/practice gap in applying ST to programme evaluation (Tikly, 2015; Cartwright, 2013). The RE framework, as Caffrey and Munro (2017) contend, explicitly emphasises the identification of 'explanations of the process of how policy works (or fails) (…) as it interacts with other systems, producing conflicts, local variation and emergent effects' (p. 466). Indeed, Westhorp (2012) has published an article in the academic journal *Evaluation*, in which she claims the existence of clear-cut commonalities and complementarities between ST and RE. Furthermore, in her view, RE is well equipped to solve the debates within the ST evaluation literature. For instance, one heated discussion

among ST-oriented evaluators lies on 'how can people rationally justify the boundaries [or the potential elements and the levels of a system] they use?' (Midgley, 2006, p. 25). Westhorp's (2012) response to that puzzle is straightforward: the rationale through which RE introduces preliminary social theories to frame the evaluation of complex social interventions settles 'which elements of the system are important for change processes [and] these are likely to be elements about which data should be collected' (p. 410).

Discussing the implications of ST principles for practice, the following excerpt from Tikly's (2015) reflective piece on the potential contributions of RE-based methodologies to comparative educational research summarises the basic logic of this approach when it comes to evaluation. In unpacking this quote, readers should note that the argument of RE researchers to study complexity in social systems is more about applying some general methodological principles for identifying the relationships between a system's different elements than it is about following unique recipes inspired, for instance, by general or fixed social theories (Mingers, 2014):

> Unlike induction, which moves from an observed phenomenon to [explanations] or deduction, which moves from general laws to hypotheses and then to observed phenomenon [a RE-based design] starts from a theory and then considers the extent to which it fits a case (…) In actual research practice, this involves assessing the explanatory ability of multiple theories (…) in the process of searching for evidence that may shed light on the contingent conditions under which a particular event (…) will occur. (Tikly, 2015, p. 247)

Therefore, RE departs from considering specific empirical events that require explanation, rather than focusing initially on causal laws to predict behaviour or study human views and perceptions without ideas of what causal forces possibly have influenced the experiences of specific social actors. In other words, 'the research team should locate the intervention under research within the dynamic policy and social systems that surround it (…) There should be an attempt to "map" complexity' (Pawson, 2013, p. 30). Such an enquiry logic resembles what qualitative researchers describe as an abductive argument (Jagosh, 2020; Tavory & Timmermans, 2014). Notably, using abduction differs from deductive reasoning – which is common, for instance, in mathematical modelling or econometric analysis – 'because we do not start with [a] general theory, but rather with the phenomena we experience' (Gibbs, 2018, p. 6) before attempting at theorising about the existence of that specific

event. Afterwards, as depicted in the second part of Tikly's (2015) quote, evaluators aim to explore multiple explanatory theories by searching (through fieldwork) for and analysing different kinds of evidence. Or as Wong (2018) puts it, 'as the evaluation progresses, data are gathered to develop [preliminary theories] and confirm, refute and refine (or "test") aspects of [them]' (p. 109).

In the remainder of this chapter, we document the experience of applying ST (via RE) in an evaluation project. One main goal of the RE evaluation framework is to assist in constructing intervention implementation models and inform the practice of policy or programme designers as they navigate complex social systems (Pawson & Tilley, 2004). We hold that successfully engaging in such an evaluation task entails 'maintaining a theoretical awareness – the systematic reflection on theory development and testing – during the entire realist inquiry' (Mukumbang et al., 2020, p. 486). This principle will remain explicit throughout our narrative of each step of our evaluation (e.g., the abduction process to define preliminary theories, sampling decisions and data analysis strategies) of a large-scale educational intervention in Colombia.

DOCUMENTING A SYSTEMS THINKING BASED EDUCATION POLICY EVALUATION

In December 2018, the Colombian National Planning Department (DNP) commissioned an assessment of the implementation process of JU – which then covered 13% of the students in public schools countrywide – in the programme's first three years of operation. Figure 9.1 is a visual representation of JU's theory of change (ToC). Because our emphasis in this chapter is on documenting our fieldwork and data analysis strategies, we emphasise the dependence of this education policy on the functioning of a multilevel decentralised governance structure.[4] Under such a scheme, public schools enrol voluntarily in the programme based on a diagnostic that the school community completes regarding their needs – for example, infrastructure, pedagogical innovation and school cafeterias – to strengthen the quality of the educational services delivered. Schools should then receive active accompaniment throughout the implementa-

---

[4]   All the study reports, including its results, are available for public consultation at the website of the governments' central policy evaluation unit: https://sinergiapp.dnp.gov.co/#Evaluaciones/EvalFin/1170

tion process of JU by the regional secretariats of education. Parallel to this, the Colombian Ministry of Education provides general implementation guidelines to all actors involved in the programme and uses the information about school needs compiled by subnational authorities to prioritise the delivery of national educational services (e.g., infrastructure plans, teaching professionalisation) across regions.
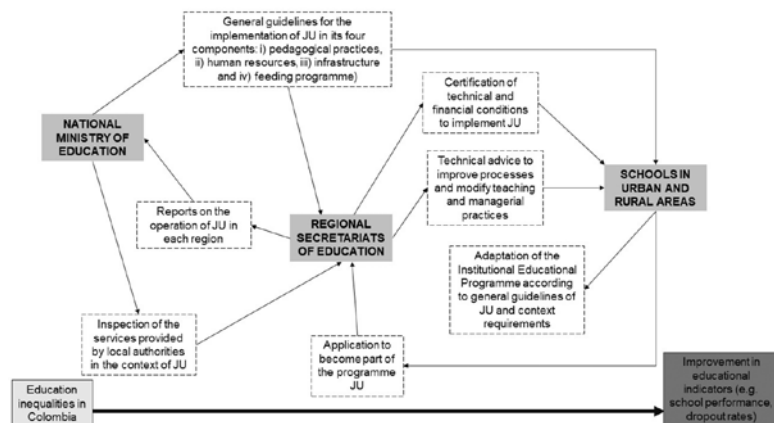
Source: Adapted from the original theory of change in the methodological reports of the project.

Figure 9.1   Jornada Unica*'s theory of change*

The remaining sections discuss the central (ST-based) aspects of our research design and its implementation to assess JU, which entailed making explicit efforts to navigate in analysing the programme's implementation at different levels (i.e., national, subnational and in local public schools), but also exploring the interactions between those levels. As necessary, we introduce some additional concepts and definitions to help us elucidate the implications of using complexity lenses to unpack different system's dynamics.

First Stage: Defining System Boundaries

As Gilmore (2019) suggests, in RE-based government-funded assessment studies, 'an intervention's Theory of Change (if available) could provide a starting framework that is already designed around notions of

how change will occur' (p. 6). Therefore, one strategy we used to begin an abductive process in our evaluation project consisted of asking the programme officials involved in different stages of the implementation of JU's ToC (in Figure 9.1) how and why they expected the intervention to work.[5] As we constantly insisted to our evaluation's audit team – from Colombia's central policy evaluation unit at the National Planning Department – evaluators are not experts in the intervention they are examining. Instead, their role consists of collecting evidence for policy planners and executioners (the real experts) for them to self-scrutinise – based on their own theories on the effectiveness of their work – their efforts to pursue specific policy-oriented goals and redirect these efforts (e.g., investments, intervention plans and targets) when required. Apart from how this argument is consistent with RE's methodology to explore complex systems (as outlined in the previous section), we would also argue that it contributed to increasing the acceptance of our evaluation's approach because it placed the voices and concerns of its final users at the centre of the study.

In Manzano's (2016) insightful analysis on the craft of interviewing in RE, she argues that these initial interactions with an intervention's stakeholders to start identifying their preliminary explanatory theories can benefit from theory-gleaning conversation techniques. In her words, '[t]he questions asked in these first set of interviews will be mainly exploratory and the wording of those questions should try to ascertain how the programme works for whom and in what circumstances' (p. 255). The queries asked for such a purpose should sound like the following: 'Is this new programme going to work for everyone? Could you explain to me the types of people and places where you think it may be more effective [and why]?' (p. 255). Mukumbang et al. (2020) also suggest that information retrieved from technical background documents (e.g., with details about the intervention's implementation procedure) serves as 'pointers to the aspects that required more probing during the interview process with the key informant[s]' (p. 495).

Based on these premises, as one of our study's first activities, we organised an expert workshop to identify underlying theories for prelimi-

---

5.   This decision is in line with the practices reported in the literature of applied ST in evaluation, which 'signals recognition of the importance of harnessing theories from within the system, rather than purely imposing theories from without' (Moore et al., 2019, p. 31).

nary enquiry that accounted for the possible outcomes derived from JU's operation in different regions of Colombia. The evaluation and audit team arrived together with a list of participants to invite to this conversation. The meeting took place at the headquarters of the Ministry of Education and consisted of two stages. First, the experts gathered in small groups defined by the implementation processes of JU, as depicted in Figure 9.1, which was built by the research team using official background documents and were asked to identify the factors that both enabled and challenged the realisation of progress in these processes. Afterwards, based on these initial reflections, the workshop leader initiated an open conversation, inviting all the participants to express their theories about how and why they expected to see specific results in JU's implementation in schools and regions. For instance, if a topic like the importance of the school's meal programmes as a determinant of the parents' involvement in school affairs was put on the table, theory-gleaning questions were posed to participants, such as the following: Why do you think that is the case? Is that situation similar in all the country's regions and with parents of students of different age groups?

The information retrieved during this workshop became the primary input for the definition of the study's preliminary theories, which accounted for the system's boundaries and indicated priority topics or themes to guide the rest of the evaluation. To minimise the excessive influence of specific social actors in the definition of the evaluation's priorities, the initial theories generated 'whilst partly emerging from discussions with stakeholders, are specified and owned more by the evaluators rather than approved and "signed up to" by the stakeholders' (Blamey & Mackenzie, 2007, p. 447). Therefore, the final definition of the preliminary theories of evaluation occurred during a post-workshop working session among the evaluation team. Finally, to express and translate those preliminary theories into researchable hypotheses, we endorsed RE's reasoning and '[its] particular approach to analysis, succinctly captured in the heuristic: context + mechanism = outcome (or C+M=O)' (Wong, 2018, p. 109). Table 9.1 presents the definitions of these items in CMO configurations, providing three examples of the eight CMOs we arrived at during this first stage of the evaluation.

The reader should note how each of these propositions attempts to make explicit use of the definition of the mechanisms (in Table 9.1) in RE and their interactions with the specific contextual features of the social systems where JU is being implemented, conditioning the emergence of its various possible outcomes. In addition, the conceptual

Table 9.1   *Examples of the initial CMOs of the evaluation of JU*

Context	Mechanism	Output
'Context describes those features of the conditions in which programmes are introduced that are relevant to the operation the programme mechanisms (…) For realism, it is axiomatic that certain contexts will be supportive to the programme theory and some will not. And this gives realist evaluation the crucial task of sorting the one from the other' (p. 7)	'Mechanisms describe what it is about programmes and interventions that bring about any effects. Mechanisms are often hidden, rather as the workings of a clock cannot be seen but drive the patterned movements of the hands (…) This *process* of how subjects interpret and act upon the intervention stratagem is known as the programme "mechanism" and it is the pivot around which realist research revolves' (p. 6, original emphasis)	'Outcome-patterns comprise the intended and unintended consequences of programmes, resulting from the activation of different mechanisms in different contexts. Realism does not rely on a single outcome measure to deliver a pass/fail verdict on a programme (...) Deciphering the reasons for such a variegated pattern can give us vital clues to the workings of programmes' (p. 8)
Examples from JU's evaluation		
CMO1	CMO 3	CMO 8
There is a general agreement among government representatives at national and subnational levels of the benefits of JU for improving the quality of education. Therefore, the main barrier to its implementation is budgetary. Thus, the local government's capacity to generate revenue internally **(C)** has a close relationship with the progress of JU in regions **(O)**. That is because governors and mayors in relatively wealthy subregions invest in the program with the firm conviction of its potential to transform education in their regions **(M)**	The clarity and the graduality in the implementation of JU are essential to gain the trust of educators **(M)** and obtain their support in the process **(O)**. That is relevant when tensions between local governments and teaching unions exist **(C)**	In contexts of poverty in the households of students **(C)**, the use of the additional class time in the teaching of arts and sports increases the interest of children and adolescents in attending school **(M)**, hence contributing to the decrease of teenage pregnancy and in reducing the consumption of alcohol and drugs among students **(O)**

Source: The definitions of the CMOs are quoted directly from Pawson and Tilley (2004).

dimensions to which the CMOs speak reflect the dynamics linking local governments and national government, local actors with subnational institutions and social interactions at schools. Although some themes in these CMOs might have been perceived as beyond the original scope of the JU programme, their exploration seemed relevant, particularly in light of Midgey's (2006) concern about the risks of having an evaluation focus that is too narrow. Moreover, an exploration of these CMO chains was necessary to engage in a comprehensive investigation of JU's casual drivers.

Second Stage: Defining Fieldwork Protocols

Emmel's (2013) book on sampling in realist-based enquiry suggests a concrete rationale for data collection planning. We construct, as he writes, 'sampling strategies through the lens of the theories we bring to the research to be tested and refined' (p. 95). From our earlier reflection on defining and investigating system boundaries, it is not hard to see how such advice resonates with the idea of studying interactions (within systems) that potentially matter (Westhorp, 2012). Such an argument about sampling also has implications for the specific role of quantitative and qualitative data sources in policy and programme evaluation. Pawson's (2013) proposal is that ideally, 'mining mechanisms requires qualitative evidence, observing outcomes is quantitative, and that canvassing contexts requires comparative and sometimes historical data' (p. 19).[6]

Our study's terms of reference had already predefined some data collection parameters. On the one hand, we had to collect perceptions about the implementation process of JU in a representative online survey of principals (N=681) from schools enrolled in the programme. Likewise, the qualitative work of the project entailed visiting ten schools in the country to conduct 30 interviews with subnational actors (public servants, principals and teachers), 20 focus groups (with parents and students) and 40 hours of non-participant observations per school.[7] Putting

---

[6]   As implied in the definitions in Table 9.1, fleshing out our mechanism in RE entails inspecting social perceptions and interactions, delving into the reflexivity process of social actors. Analysing such complexity aligns more with the ethos of qualitative instead of quantitative research (Gibbs, 2018).

[7]   We also consulted with national-level stakeholders. One central activity to collect such data was the workshop with experts we mentioned before to define the study's preliminary theories.

RE's fieldwork planning principles (as outlined in the last paragraph) in practice in our assessment entailed different considerations. For instance, for selecting the sites to visit, an agreement was made with the audit team to inspect official data sets from various governmental sources to search for any comparative indicators of municipalities and schools that could serve as a proxy for measuring the contextual features defined in our preliminary theories. We also agreed that we would use the quantitative data from the survey to search for some proxy measures of the outputs in our CMOs and, finally, that the study of mechanisms would rely on the information retrieved from the interviews, focus groups and field observations.

Table 9.2 illustrates the results of our sampling exercise (with information relevant only for the case of the exemplified CMOs in Table 9.1). On the one hand, we considered the first set of indicators, from A to D, to ensure we visited schools with enough exposure to the programme – in terms of implementation time and number of students in attendance – at rural and urban sites. On the other hand, we searched for statistical information relevant to studying CMOs in specific contexts where those same CMOs indicated it was more likely to assess particular mechanisms at work. For example, CMO1 stated that one might expect to see more progress in the overall implementation of JU in subregions with relatively good fiscal capacity than in those depending on funding from the central government. Thus, using secondary information, we computed indicator F in Table 9.2, which ranked the different municipalities in the country from one to six, where the latter value represented high fiscal independence. We agreed with the audit team that we would include at least two cases of schools located in municipalities with values five or six in this indicator and two with values of either one or two. The same strategy was applied to indicator E to define sites relevant to the study of CMO8 (Table 9.1). We repeated this exercise with the available information concerning contextual characteristics on the other CMOs of our study, ultimately defining 10 sites for our fieldwork.[8]

---

[8]   Notably, we did not find any statistical data to measure the trust levels between teachers and local governments to address CMO3's contextual description, so we could not replicate the same exercise with information relevant to this preliminary theory. We mention this last challenge here to make the point that it is possible to encounter this kind of situation in RE-based designs. As Pawson (2013) reminds us, CMOs are not mathematical formulas or catalogues for a simple 'tick-box format' data collection exercise: 'The immediate priorities

Table 9.2   *Results of the (qualitative) sampling exercise*

Colombian Subregion/Indicator	School-level characteristics					Municipal-level characteristics
	Time in JU	Students in JU	% Students in JU	Rural school	Socioeconomic status (SES) of students	Level of financial dependency
	A	B	C	D	E	F
Caribe	> 2 years	464	100	No	6	2
Caribe	> 2 years	1,196	100	No	1	6
Cento Sur-Amazonia	> 2 years	290	100	Yes	2	6
Centro Oriente y Bogotá DC	2 years	998	100	No	5	5
Centro Oriente y Bogotá DC	> 2 years	412	100	Yes	5	5
Eje Cafetero & Antioquia	> 2 years	572	42	No	6	1
Llanos	2 years	81	100	Yes	2	4
Pacífico	> 2 years	390	100	Yes	2	6
Pacífico	> 2 years	594	100	No	6	2
Pacífico	> 2 years	88	100	Yes	1	4

Source: Adapted from the original table in the final report of the project.

Next, in this discussion about fieldwork technicalities, we also reflected on how to study causal mechanisms using qualitative sources, here following Pawson's (2013) advice. The RE literature suggests realist – or theory-based – interviews suited for this purpose. At the core of this conversational approach is the idea that instead of simply collecting perceptions of people, for example, in an open-ended interview fashion, '[the]

of evaluation research are to respond to the research brief, to deal with the given substantive issue, and to contribute to policy development – rather than to aim for methodological purity' (p. 27).

subject matter of the interview is the researcher's theory and interviewees are there to confirm, falsify and, basically, refine the theory' (Manzano, 2016, p. 344). Manzano (2016) further explains the following:

> This relationship – described as a teacher–learner cycle – is distinctive of realist evaluations. It starts with teaching the interviewee 'the particular programme theory under test' and then 'the respondent, having learned the theory under test, is able to teach the evaluator about those components of a programme in a particularly informed way' (Pawson and Tilley, 2004: 12). Therefore, the roles of teacher and learner are not static but become interchangeable between the interviewer and the interviewee during the process of thinking through the complexities of the programme. (p. 344)

Therefore, we designed our interview protocols to explicitly motivate teacher-learner conversation dynamics between the field researchers and different JU's stakeholders. In that vein, after some general opening questions to help interviewees refresh their memory and activate their reflexive thinking about the programme, we introduced what we called provocative statements in the interview protocol. Instead of questions, these consisted of phrases that announced – using a nonformal language – the evaluation's initial theories (or CMOs). Table 9.3 shows two examples of these statements associated with CMO1 and CMO8. We expected the interviewees to initially show agreement or disagreement with the different statements and then justify their answer by conveying their interpretations (or theories) of the underlying mechanisms that account for the specific outcomes emerging from JU's specific implementation processes and activities. The interactions we experienced during the data collection process suggested that this was an appropriate strategy to motivate the reflections of stakeholders and beneficiaries of the programme concerning the factors or circumstances that could potentially explain its different outcomes.

Finally, the second column in Table 9.3 links the statement to specific social actors in the fieldwork plan. Adding to Emmel's (2013) reasoning about sampling in RE, Manzano (2016) clarifies that, under this worldview, 'the importance is not on "how many" people we talk to but on "who", "why" and "how"' (p. 349). From the methodological reflection presented so far, it is reasonable to suggest that different people (from policy designers to final beneficiaries of a programme) can be key informants of a specific system's dynamics, here depending on their experience and position within the system. Based on this premise, we agreed with the evaluation's audit team that certain social actors were

Table 9.3   *Examples of provocative statements in the interview*
              protocols

To explore the mechanism in CMO 1	Local education authorities
Without a doubt, JU enables accomplishing national and regional educational goals. If local governments have enough revenues, there is no reason not to see progress in the implementation of the programme at the school level.	School principals
To explore the mechanism in CMO 8	School directors
	School teachers
The JU programme helps improve school performance and reduces the exposure of students to risks such as teenage pregnancy and the consumption of alcohol and drugs only if the additional school time afforded by the programme is invested in vocational training, sports and arts.	Focus groups with parents
	Focus groups with students
	Non-participant observations

better positioned to discuss specific CMOs. As an additional note, we argue that these kinds of practical decisions – grounded in the epistemological foundations expounded so far – are essential in evaluation studies in settings with considerable time constraints to collect and analyse massive amounts of qualitative information.

Third Stage: Data Analysis and the Presentation of Results

Table 9.4 is an adapted version of the template we used in the data analysis stage of the evaluation, where we combined the building of single case study reports (for each school we visited) with a subsequent cross-case analysis. The first step in this strategy was to read the transcripts of the interviews, focus group discussions and fieldnotes (from the non-participant observations), extracting the main points and quotes alluding to each of our provocative statements.[9] This exercise invited evaluators to first reflect on evidence from each regional visit to interro-

---

[9]   Note that in the format, we maintained the logic of linking the discussion of each statement with information emerging from the specific conversations with the actors we assumed were key informants for each topic. Because that strategy posed the risk of ignoring discussions with people we did not anticipate in our exercise illustrated in Table 9.3, we added a section to document qualitative information that the analysts found hard to document in the tables.

Table 9.4   *Example of a part of the template for each case study report*

STATEMENT: **Without a doubt, JU enables accomplishing national and regional educational goals. If local governments have enough revenues, there is no reason not to see progress in the implementation of the programme at the school level.**

Field notes from non-participant observations	
Relevant findings to problematise the statement	Main support quotes (max 10)

Interview with representatives of the SEC			
Key arguments that refer explicitly to the statement	Relevant support quotes (max 5)	Key arguments from other sections of the interview	Relevant support quotes (max 5)

Interview with the school director			
Key arguments that refer explicitly to the statement	Relevant support quotes (max 5)	Key arguments from other sections of the interview	Relevant support quotes (max 5)

In this space, you can add notes and general reflections of the different sources of qualitative information that you feel do not respond to any of the categories in the previous tables.

gate the original CMOs of the project. In the second phase, the leading qualitative researcher of the team made a complete reading of all the single case study reports using NVivo software and some of its data analysis tools (like attributes and coding matrices). This step in the analysis allowed for the emergence of new insights (in the form of codes or categories arrived at inductively) related to social dynamics not necessarily anticipated in the preliminary explanatory theories. The strategy outlined above recreates the logic of a realist-informed qualitative content analyst discussed in Parra et al. (2021).

At this point, it is worth emphasising that we incorporated computer software in our analysis, as some RE practitioners recommend (Dalkin

et al., 2021) but only after making an initial reading of the raw data with the preliminary CMOs of the evaluation in mind. In our experience, computer software can assist in analysing complexity only if researchers make explicit efforts to avoid the temptation of splitting information into isolated categories (e.g., using nodes), which tends to remove context specificities within which conversations took place. Hence, our strategy invited a first reading and contrasting of multiple sources vis-à-vis a consideration of the potential system's boundaries (expressed in the form of CMOs). At the same time, this strategy allowed for the emergence of new categories in the analysis of qualitative data. This is an important step, especially given the emphasis of RE on going beyond theory verification or refutation to theory refinement based on new insights from empirical research.

Before presenting our final report, we organised (in collaboration and with approval from the evaluation's audit team) one last workshop with education experts in the country to discuss our ideas for policy recommendations. In this final evaluation stage, we did our best to balance the potential evaluation users' expectations about our work (e.g., to provide practical insights on how to improve JU) with our theoretical interest in maintaining consistency with general ST principles. Pawson and Tilley's (2004) view about the possibility for REs to prescribe good policy advice was illuminating in helping us navigate this tension:

> The school of theory-based evaluation, of which realist evaluation is an affiliate, has always described its appointed task as 'enlightenment' as opposed to 'political arithmetic' (…) Perhaps the best metaphor for the end-product is to imagine the research process as producing a sort of 'highway-code' to programme building, alerting policymakers to the problems that they might expect to confront and some of the safest measures to deal with these issues. An evaluation highway-code could never provide the level of prescription or proscription achieved in the real thing, the point of the parallel being that the highway-code does not tell you how to drive but how to survive the journey by knowing when, where and for what to keep your eyes peeled. (pp. 20–21, original emphasis)

Returning to the discussion in section two about uncertainty as one essential consideration in ST-based enquiry, we acknowledge that policy evaluators cannot predict (at least not with statistical precision) what will happen in a system after adjusting their (policy) interventions. Therefore, instead of including statements such as *you should do this*, we preferred to point out *what might happen if you continue doing that*. For instance,

our study reported some problems with the coordination between national and subnational authorities in delivering educational services to schools and accompanying their in-house planning processes. Consequently, instead of suggesting specific recipes to solve those coordination challenges, our message to the end users of the evaluation highlighted that such a situation might keep happening unless decisions were made to revise the decentralised institutional arrangements (between authorities) under which JU operates.

We perceive a positive reception of such an approach by the audit team and the experts in the workshop. It also invited interesting discussions about JU as a subsystem within a national education system. This is relevant because, in our experience, high-level policy discussions in a country like Colombia often remain at the micro level of educators' responsibilities or the definition of, for instance, protocols for how to use technology and different pedagogical strategies in the classroom (Parra, 2018). However, to our surprise, we saw more in-depth systems-informed conversations emerging about the underlying mechanisms behind the expected results of JU across the country involving, for instance, educational decentralisation in Colombia or the meaning of quality in education for different groups or members of society. We believe that the evaluation methodology described in this chapter enabled and invited those kinds of conversations addressing more structural drivers of education policy results. Altogether, these outcomes justify the enthusiasm of authors like Mason (2009) regarding the potential of ST thinking to inform sustainable, positive, system-wide change and development in education.

CONCLUSION

Throughout this chapter, we have argued that the RE approach represents a promising methodological pathway for operationalising ST in policy-focused educational research. To demonstrate this potential, we reflected on our experience applying RE tenets to evaluate the implementation process of a large-scale educational intervention in Colombia – the JU programme. To this end, one of the principal arguments advanced here is that studying complexity in social systems requires a sustained effort to integrate ST-based principles in the different stages of the research process, specifically while collecting and analysing (qualitative and quantitative) data. We also emphasised the importance of avoiding the temptation to jump directly from research findings to specific policy

prescriptions. Ultimately, as a result of our approach, we did not only report insights that depict the workings of the assessed programme – and the mechanisms on which it is based. We argue that it was also able to engender meaningful and broad discussions among policymakers going beyond technical discussions of isolated intervention components bringing in the interactions of multiple levels and elements of the Colombian education system.

REFERENCES

Blamey, A. & Mackenzie, M. (2007). Theories of change and realistic evaluation: Peas in a pod or apples and oranges? *Evaluation*, 13(4), 439–455.

Broad, R. (2006). Research, knowledge, and the art of 'paradigm maintenance': The World Bank's Development Economics Vice-Presidency (DEC). *Review of International Political Economy*, 13(3), 387–419.

Byrne, D. (2005). Complexity, configurations and cases. *Theory, Culture & Society*, 22(5), 95–111.

Caffrey, L. & Munro, E. (2017). A systems approach to policy evaluation. *Evaluation*, 23(4), 463–478.

Cartwright, N. (2013). Knowing what we are talking about: Why evidence doesn't always travel. *Evidence & Policy*, 9(1), 97–112.

Dalkin, S., 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(49), 2–7.

Dalkin, S., Forster, N., Hodgson, P., Lhussier, M. & Carr, S. (2021). Using computer assisted qualitative data analysis software (CAQDAS; NVivo) to assist in the complex process of realist theory generation, refinement and testing. *International Journal of Social Research Methodology*, 24(1), 123–134.

Edwards, B. (2018). *Global education policy, impact evaluations, and alternatives.* Palgrave Macmillan.

Edwards, B., Morrison, J. & Hall, S. (2020). The suspect statistics of best practices: A triple critique of knowledge production and mobilisation in the global education policy field. *Globalisation, Societies and Education*, 18(2), 125–148.

Emmel, N. (2013). *Sampling and choosing cases in qualitative research. A realist approach*. Sage.

Gibbs, G. (2018). *Analysing qualitative data* (2nd edn). Sage.

Gillies, J. (2010). *The power of persistence. Education system reform and aid effectiveness. Case studies in long-term education reform*. United States Agency for International Development (USAID).

Gilmore, B. (2019). Realist evaluations in low- and middle-income countries: Reflections and recommendations from the experiences of a foreign researcher. *BMJ Global Health*, 4, 1–8.

Jagosh, J. (2020). Retroductive theorising in Pawson and Tilley's applied scientific realism. *Journal of Critical Realism*, 19(2), 121–130.

Joyce, K. & Cartwright, N. (2020). Bridging the gap between research and practice: Predicting what will work locally. *American Educational Research Journal*, 57(3), 1045–1082.

Manzano, A. (2016). The craft of interviewing in realist evaluation. *Evaluation*, 22(3), 342–360.

Mason, M. (2009). Making educational development and change sustainable: Insights from complexity theory. *International Journal of Educational Development*, 29(2), 117–124.

McGill, E., Er, V., Penney, T., Egan, M., White, M., Meier, P., Whitehead, M., Lock, K., Anderson de Cuevas, R., Smith, R., Savona, N., Rutter, H., Marks, D., de Vocht, F., Cummins, S., Popay, J. & Petticrew, M. (2021). Evaluation of public health interventions from a complex systems perspective: A research methods review. *Social Science & Medicine*, 272, 1–18.

Midgley, G. (2006). Systems thinking for evaluation. In B. Williams & I. Imam (Eds), *Systems concepts in evaluation. An expert anthology* (pp. 11–34). American Association of Evaluation.

Mingers, J. (2014). *Systems thinking, critical realism and philosophy. A confluence of ideas.* Routledge.

Moore, G., Evans, R., Hawkins, J., Littlecott, H., Melendez-Torres, G., Bonell, C. & Murphy, S. (2019). From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation. *Evaluation*, 25(1), 23–45.

Mukumbang, F., Marchal, B., Van Belle, S. & van Wyk, B. (2020). Using the realist interview approach to maintain theoretical awareness in realist studies. *Qualitative Research*, 20(4), 485–515.

Parra, J. D. (2018). Critical realism and school effectiveness research in Colombia: The difference it should make. *The British Journal of Sociology of Education*, 39(1), 107–125.

Parra, J. D., Said-Hung, E. & Montoya-Vargas, J. (2021). (Re)introducing critical realism as a paradigm to inform qualitative content analysis in causal educational research. *International Journal of Qualitative Studies in Education*, 34(2), 168–182.

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

Pawson, R. & Tilley, N. (2004). *Realistic evaluation*. British Cabinet Office.

Radinger, T., Echazarra, A., Valenzuela, J. P. & Guerrero, G. (2018). *OECD reviews of school resources: Colombia 2018*. OECD Publishing.

Spivack, M. (2021). Applying systems thinking to education: The RISE systems framework. Research on Improving Systems of Education (RISE) Programme.

Tavory, I. & Timmermans, S. (2014). *Abductive analysis. Theorising qualitative research*. The University of Chicago Press.

Tikly, L. (2015). What works, for whom, and in what circumstances? Towards a critical realist understanding of learning in international and comparative education. *International Journal of Educational Development*, 40, 237–249.

Turner, J. & Baker, R. (2019). Complexity theory: An overview with potential applications for the social sciences. *Systems*, 7(1), 1–22.

Westhorp, G. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation*, 18(4), 405–420.

Wong, G. (2018). Getting to grips with context and complexity − The case for realist approaches. *Gaceta Sanitaria*, 32(2), 109–110.

10. Can systems thinking tools help us better understand education problems and design appropriate support? Reflections on a test case

Elena Walls and Laura Savage

This chapter recounts the experience of piloting an 'education system diagnostic' using systems thinking tools that have been well established in other sectors but minimally used in education. The authors conceived of the idea and were involved throughout the two-year period of designing the pilot and seeing it through. One of the authors took the pilot further into a second-country case study. We then worked as education advisers for the US and UK government aid agencies and learned much from this effort.[1] In this chapter, we discuss (1) why we wanted to do it, (2) what happened and (3) what we have taken from this into our work.

We sought to test the hypothesis that a system diagnostic in the education sector could improve our understanding of how education systems function and, in turn, how this could lead to different ways of working with education systems to deliver improved learning outcomes. We first learned that 'system diagnostic' is a very much misunderstood phrase. The word 'system' can be a catch-all, as the other chapters in this book point out. The word 'diagnostic' immediately draws out analogies to

---

[1]   Elena Walls is Senior Monitoring, Evaluation and Learning Advisor for the US Agency for International Development (USAID). At the time of this pilot, Laura Savage was a Senior Education Adviser for the UK Department for International Development (DFID), which became the Foreign, Commonwealth and Development Office (FCDO) in August 2020. Laura is now Executive Director of the International Education Funders Group. Tjip Walker at USAID and Matt Clancy at DFID/FCDO were also core team members, though are not authors here. The views expressed here are those of the authors and not of their respective organisations.

medical practice and expectations of a formulaic exercise. Our journey to pilot a system diagnostic in education was one in which we unpacked the meaning of 'system' and tried to unpack others' definitions by setting common boundaries around the education system we were working within. We revised – and revised again – our approach of getting to a point of understanding the dynamics within this education system. We came to realise that even the idea of a 'diagnostic' could be anathema to systems thinking, prompting assumptions of solutions and quick fixes. Our initial expectation that a system diagnostic could be a tangible way of introducing systems thinking into educational conversations was robustly challenged. We learned that system diagnostics would not be welcomed everywhere, that they would raise more questions than could be answered and that they would pose unpalatable challenges to the status quo. We built a catalogue of examples of how *not* to do a system diagnostic. However, we ultimately found that systems thinking tools *can* help draw out new perspectives on an education community's problems. To us, this points to the promise and potential of systems thinking as a framework for doing *education* development differently.

USAID/FCDO PARTNERSHIP TO TEST A NEW APPROACH

In mid-2018, we were part of a small network of individuals within the global education aid architecture who were convinced that collective problem identification at the local[2] level could help. To some extent, this interest in 'diagnostics' was driven by the growing recognition of 'the learning crisis': that investments in the education sector have increased school enrolments but not led to improvements in educational outcomes. We will not further recount here the reasons for our frustration and despair; the education statistics are all in this book's introduction. The 'problem' to be tackled was that there was a limited understanding of the reasons for why education outcomes remained so low, despite the investments made.

We identified with a small but increasingly influential school of thought in development theory that called for 'doing development differently' (Wild et al., 2016; Teskey, 2017). In particular, though not

---

2   Local, to us, meant country, district or school – wherever the locus of power and decision making lay.

knowingly at the time, we intuitively agreed with the 'problem-driven iterative adaptation' approach (Andrews et al., 2015). This (and it deserves quoting at some length) is as follows:

> [It is] about building capability through the process of solving good problems. It's not about finding *the* solution and then replicating that solution; it places emphasis on the *process* of solving problems, not the solutions … It is not easy, or without real risk but ultimately it is a more sustainable approach because it infuses legitimacy into change processes. (Andrews et al., 2015, p. 3)

This sector-agnostic theory aligned with our own experiences working in country on education reform. We believed that collective problem identification – a process of gathering education stakeholders together to dig deep into analysis of education data and why the outcomes were not improving – would result in greater coherence of response, stronger shared accountability for progress and less cut-and-paste responses to assumed problems.

So why bring systems thinking tools into this problem identification process? The advocates of system thinking argue that developing a better understanding of the education sector through 'system diagnostics' will help target interventions at key challenges and increase the likelihood that these interventions will lead to sustained improvements in educational outcomes. System diagnostics seemed to promise the potential of an answer to the question of 'why': Why, despite the investments made, were student learning levels remaining persistently low? Why could evidence point to what did not work far more than what did? Yet as pointed out in a DFID-commissioned study (Faul, 2016), no existing tool in global education truly served as a 'system diagnostic' that could seek to understand *how* a particular system (whether defined as national or by closer geographical or other boundaries) works and *why*.

Our collaboration in applying systems thinking tools was one of several 'problem identification' exercises being developed and tested among the global education aid community in 2018–2019. For some, a 'diagnostic'[3] amounts to the descriptions of symptoms, whether in the form of a narra-

---

[3]   It is useful to distinguish between a 'diagnostic of a system' and 'diagnostics used within the system'. Several of the latter exist in education systems, for example, the inspection frameworks that are used to assess the quality of schools or school report cards that are used in conjunction with school improvement grants. The focus of this collaboration is on the diagnosis of a system.

tive (such as sector reviews or reports that document policy intent only) or a collection of statistics, such as the Global Partnership for Education's Results Framework (GPE, 2019) or the World Bank Education Policy Dashboard (World Bank, 2019). However, in the medical sense, a diagnosis is more than descriptive data. A diagnosis is the process of determining which disease or condition explains a person's symptoms. Thus, although data and descriptions serve a useful purpose, a diagnosis needs to go further, providing a coherent explanation of *why* those symptoms exist.

A second difference between our project and others revolves around the use of the term 'system'. For some, the term is used as a synonym for 'sector'. We (and others) use the term in a more technical sense. A system refers to a set of parts that perform a collective function. When applied to education, this means the educational activities, institutions and stakeholders within a given context (whether global, cross-national, national or local). The education system refers to the interconnected sets of actors (governments, civil society, the private sector, universities, individual citizens, and others) that jointly contribute to the outcomes and activities related to education within a given context. This definition has implications. No one actor or input has the ability to affect learning outcomes because the collective outcome depends on the ways the actors in that system interact with one another. A 'system diagnostic' is a holistic assessment of a system – of which the boundaries defined at the outset – that provides a view of the inputs and resources, relationships within the system, the system's intended and actual functions and the political economy within the system.[4] Thus, a 'system diagnostic' differs from a 'sector assessment', which is a more commonly used term, in that it (a) does not seek to 'assess' (but to diagnose), (b) deliberately seeks understanding of the dynamics, relationships and complexity (whereas a sector assessment might not) and (c) strives for mindset and behaviour change through the process rather than seeking the descriptive justification for a course.

A third difference relates to the *use* of a system diagnostic tool. For some, it is enough to *understand* a system as it is: its dynamics and how those dynamics account for the current levels of performance. For us, as actors within operational agencies, a diagnostic is needed not only to capture the system as it is, but also to provide the information needed for

---

4   See also DFID Education Policy 2018 (DFID, 2018).

identifying how we could work with our local partners to produce the education system we all want it to be. Thus, a diagnostic cannot just be an explanation. It must also serve as the basis for *action*.

Yet even for those who wish to use system diagnostics as the basis for designing an intervention or support for a system, there are different ideas about the process to undertake – and, therefore, the information a diagnostic needs to provide. One approach suggests that effective interventions will need to fulfil three criteria: technical effectiveness, political feasibility, and operational practicality (Harris et al., 2013). In other words, through examining the context and examination of possible interventions, the aim is to identify reforms that have been shown to be effective, garner support from the government and other key stakeholders and find what can be carried out by those in charge of implementing them. For those in this camp, it is technical effectiveness that is the most important. Thus, a diagnostic needs to reveal where technical deficiencies lie.

In this collaboration, we intended to take a different approach. Accepting that technical effectiveness, political feasibility and operational practicality are the key characteristics of successful interventions, we planned to start by examining the political and operational factors. Thus, we sought a diagnostic approach that would not only capture systems dynamics and current performance, but also elicit the leverage points in the system: the places where the prospects for initiating positive change would seem the most promising. Once the leverage points were identified, we wanted to consider which technical options would be the most effective in that context to activate change. Again, the search was for the technical, political and operational sweet spot. However, the process of getting there was quite different.

WHAT HAPPENED?

The initial plan was to develop and pilot an education system diagnostic in two countries by September 2019, here by working with USAID and DFID in-country teams. First, we designed a three-step tool. We would use systems thinking tools as the core part of this exercise, such as those used successfully by USAID contractors in other sectors in the past. These included causal loop mapping, which helps to visualise how different variables in a system are perceived to be causally related. The Strategic Program for Analyzing Complexity and Evaluating Systems

(SPACES)[5] team had used this across different sectors and geographies and was one team selected. Another model we liked was FHI360's Whole-System-in-the-Room mapping tool, which helps to define known issues and challenges around a system based on research and consultations. However, we thought that leaping straight into these tools would be too ambitious, so we planned a first step of data mapping: providing technical support to the government to map existing data, including as many proxies of system performance as possible. The aim of this step was to open the conversation by telling the story of education statistics in that particular context and to get to the point at which it became obvious that existing data could only tell us so much about why learning outcomes were so low. We also added a third and closing step: applying USAIDs existing 5Rs systems thinking framework (USAID, 2016). This would help bring the collaborative problem identification achieved in step two into a digestible format for action.[6]

We hoped that these three steps – (1) data visualisation, (2) collaborative problem identification using systems thinking tools and (3) post hoc mapping to a systems framework – would form a coherent system diagnostic. This three-step tool embraces the inherent complexity of an education system by capturing its unique dynamics and appreciating how those dynamics can influence the effects of interventions, including those that have been successful in other contexts. The tool was designed with the inherent appreciation that introducing change to complex systems usually requires more than a single intervention and entails several interventions working in concert to activate several points of leverage. This tool would enable us to seek perspectives and inputs from a wide range of stakeholders at the national and local levels across public and private actors. This was important not only for assembling a holistic understanding of the system, but also for building support for reform. The tool would produce maps and other visual depictions of the system

---

[5]   SPACES is a contract under the broader MERLIN mechanism funded by the USAID Global Development Lab; it stands for the Strategic Program for Analyzing Complexity and Evaluating Systems. The implementer for the Rwanda diagnostic was John Hopkins University in partnership with LINK.

[6]   The definition of 'system' in the current study is aligned with USAID's Local Systems: A Framework for Supporting Sustained Development policy document and USAID's The 5Rs Framework in the Program Cycle. For more information, see SPACES MERL Systems and Complexity White Paper: https://linclocal.org/wp-content/uploads/2017/02/pa00m7qz.pdf

that could promote a shared understanding of the education system and, thus, a stronger basis for policy- and decision making, as well as more coordinated efforts by local and international actors to introduce reform. It would also provide most of the information to frame a strong theory of change for donors.

The next step was to find countries in which to work. This took well over a year. We had hoped to try this in two countries. We looked for places where the governments were at a stage in their reform cycle that they wanted to drill into why learning outcomes were low and design new strategies in response. We approached countries via our colleagues inside those countries who held relationships with the governments. Our colleagues were cautious – about trying new tools, about the timing, about the optics of this being externally led or about the plethora of analytical tasks already being funded across the donor group. These were all very valid concerns. We identified two countries interested and completed step one – the data analysis. However, both countries pulled out.

In the end, we (FCDO and USAID) piloted this in Rwanda together and USAID ran the process again in the Democratic Republic of Congo (DRC) some months later. In both cases, the diagnostic was implemented by a research group commissioned by USAID (SPACES in Rwanda and the Monitoring, Evaluation and Coordination Contract (MECC) in DRC).[7] The information produced through the diagnostic was used by the USAID project design staff. DFID funded a process evaluation of the Rwanda pilot that was completed by Oxford Policy Management.

Rwanda

The research studies implemented in Rwanda between 2012 and 2019 found a range of barriers to improving learning outcomes, including poverty and other socioeconomic factors, as well as education sector-specific factors like overcrowding and poor teacher pay (Laterite, 2018; World Bank, 2018). During the first data phase, SPACES completed secondary desk analysis to build an initial causal loop diagram of

---

[7]   The Strategic Program for Analyzing Complexity and Evaluating Systems (SPACES), implemented by Johns Hopkins University, and Monitoring, Evaluation, and Coordination Contract (MECC), implemented by International Business & Technical Consultants, Inc. (IBTCI). Both mechanisms were funded by USAID.

the Rwanda education system.[8] Combining this with stakeholder consultations, the SPACES team identified the following two key questions as critical for improving our understanding of the key issues and barriers to primary grade teaching and learning in Rwanda:

1. Student experience: why do some students dislike school?
2. Teacher experience: why are some teachers unable to fully utilise their skills, knowledge and expertise in the classroom?

SPACES then facilitated five half-day workshops in two districts (one rural, Kirehe, and one urban, Kicukuro) and one national workshop in Kigali in late 2019. In each district, one workshop was conducted with school-level stakeholders (parents, teachers and headteachers) and one with government stakeholders (village leaders, sector/district education officials, and development and implementing partners). This split was a methodological decision based on analysis of what would enable the most valuable discussion. The primary goal of the workshops was to generate new insights into challenges in the education system that could lead to improvements in Kinyarwanda literacy among primary school students (primary 1 to 3).[9] Through participatory workshops, SPACES aimed to uncover deeper insights into the elements, connections and possible leverage points in this system that are less clear in the literature on Rwanda education.

The specific systems thinking tools for step two used by SPACES in Rwanda were influence mapping and system diagramming. These methods are situated within a family of systems mapping methods. Influence mapping belongs to the set of systems mapping methods that are designed to generate insights about a particular problem that different stakeholders might have different perspectives about. It is not designed to determine causality but rather to improve our understanding of the perspectives that different stakeholders have that influence the key outcomes of interest. The second method – system diagramming – is designed to portray a system in its complexity and interrogate the relationships between different elements of the system.

---

[8] This was a divergence from the first step of data analysis and presentation that had been planned because the contracted time for more 'traditional' analysis had by now run out with the two false-start cases.

[9] Kinyarwanda is the national language.

Direct observations of three out of the five workshops showed a high level of engagement. The participants visibly enjoyed the process and commented on the value of the discovery process when they were building out the influence map from the first level of the five key influencers to subsequent levels, arriving at unexpected influencers that were then found to be connected to other influencers. Comparing perspectives (e.g., the 'parent perspective' table visiting the 'teacher perspective' table and learning about teacher perspective) helped the participants to gain a better understanding of the constraints faced by other stakeholders. Additional evidence of the value of the workshop for the participants was the observed fact that most of the participants took photos of the created maps at the end of the workshop and commented on how they would like to replicate the process in their respective communities or organisations.

Based on the data from the workshops, stakeholder consultations and the literature review, the SPACES team constructed a system diagram of the components and processes that affect a Rwandan child's literacy. The system diagram shows a number of interrelated components of various degrees of influence on school enrolment, attendance and learning outcomes, including those factors related to family, school and classroom, teachers, communities and government. The diagram included four major 'outcomes' nodes ('my initial enrolment in school', 'my regular attendance in school', 'my comprehension of the material' and 'my literacy in Kinyarwanda language') and five areas of influence on the outcomes: family, community, school and classroom, teacher, and the government. Within those areas, the diagram listed 153 distinct factors that influence the outcomes. Overall, the diagram conveyed the complexity of the primary education system which impacts the four outcome areas.

The diagnostic study produced a range of recommendations for USAID and for the Government of Rwanda, including supporting children's early life and nutrition, providing socioemotional support to teachers and students, changes in disciplinary measures, simplifying early grade curriculum, strengthening relationships among the learners, their families, teachers, schools and communities, and providing families and communities with the necessary resources to support school attendance and learning (SPACES, 2019, p. 16).

USAID DRC Diagnostic Study

The systems diagnostic study implemented in the DRC was part of a larger needs assessment that focused on the following questions:

1. What impact may present policies/practices at the local/national levels have on the literacy learning and reading skill acquisition of Batwa children, girls, and children with disabilities in North Kivu and South Kivu?
2. What do formal/informal learning environments (where literacy learning occurs) look like for Batwa children, girls, and children with disabilities in North Kivu and South Kivu?
3. What are the common knowledge, beliefs and practices around literacy and language among Batwa children, girls, and children with disabilities in North Kivu and South Kivu?

Similar to the objectives of the study in Rwanda, the DRC diagnostic sought to identify the opportunities and challenges in literacy acquisition. In contrast to the study in Rwanda, the DRC assessment focused on specific population subgroups: Batwa children, girls, and children with disabilities in the eastern DRC. The assessment used desk research followed by key informant interviews (KIIs) and focus group discussions (FGDs) methods with education officials at the provincial and local levels and community leaders, parents, school directors, teachers, students, and NGOs based in North Kivu, South Kivu and Kinshasa. The systems diagnostic portion of the assessment used the influence mapping approach and was held after the KII and FGD data collection and analysis. The participants in the workshops included education officials, school directors, teachers, community leaders and parents. The workshops focused on the factors that limit educational access and literacy acquisition of Batwa children, girls, and students with disabilities. From these discussions, an initial 'systems map' was developed to see how the problem's moving parts are connected and identify the key leverage points that will allow USAID to develop effective interventions.

The implementation of the assessment in the summer and fall of 2020 was hampered by COVID-19 pandemic-related restrictions on mass gatherings and travel. System diagnostic workshops were limited to 15 participants. In total, the MECC team conducted four half-day workshops in Goma (North Kivu) and four half-day workshops in Bukavu (South Kivu), totalling 119 participants between the two. In each province, four

diagnostic workshops were organised with different stakeholder groups: (1) education officials, school heads and teachers; (2) parents and local organisations supporting children with disabilities; (3) parents and local leaders supporting Batwa children; and (4) parents and local leaders supporting girls.

Following the influence mapping approach, the participants were asked to identify the root causes of the key factors that impact the ability of the in-focus populations to learn to read in primary school. The inter-connected problems and causes were then mapped out into a multiple cause diagram following these steps: (1) The organisers presented infor-mation to the participants so that they clearly understood the goal of the workshop and what prior knowledge was available. (2) The participants discussed the central problem to be solved, thus ensuring a common understanding of the problem statement. (3) The central problem was written for all to see on a large piece of paper. In small groups, the par-ticipants were asked to suggest various reasons to explain an identified problem. The note-taker recorded the factors and causes by placing each suggestion on a smaller piece of paper around the central problem. This three-step process was repeated until the participants exhausted their ideas. It was expected that some causes or reasons would appear in multiple places on the diagram. The participants moved the causes and reasons around until they formed a reasonable representation of how the interlocking causal chains are connected, which most people in the small group seemed to accept.

The resulting system map helped build a common understanding of the system's complexity and show how causes and effects relate to each other (IBTCI, 2020, pp. 6–7). The study report indicates key issues identified in the system diagnostic workshops that relate to low learning among Batwa children. Poverty, culture and lifestyles, and discrimina-tion were the most central factors identified in the systems diagnostic workshop for why Batwa students are left behind in literacy instruction. The causes behind the relevance of these factors varied somewhat by stakeholder group, yet highlighted a tension between large, centralised government reasons for discrimination and poverty and hyper-localised ways that poverty and discrimination impact Batwa's educational trajec-tory. In all workshops, poverty and discrimination were directly linked to poor literacy acquisition while also acting as a cause behind other factors, revealing the complex connections among poverty, discrimination and other factors that impede access to high-quality literacy instruction (IBTCI, 2020, p. 72).

The study team produced multicause factor diagram maps for each of the study populations (Batwa children, children with disabilities and girls) and for each stakeholder involved in the assessment (education authorities, NGOs, community leaders and parents). Figure 10.1 shows the system map for Batwa students from the point of view of community leaders (IBTCI, 2020, p. 75).
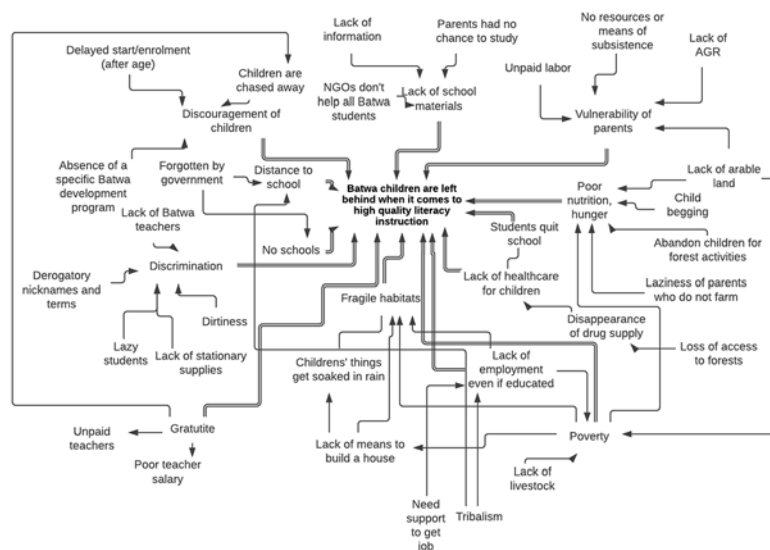
Source: IBTCI (2020).

Figure 10.1   Multicause factor diagram for Batwa students

The study design which allowed for the comparison and contrasting of different stakeholder opinions regarding the factors contributing to poor learning among Batwa children was especially helpful. For example, in contrast to the education officials and NGOs indicating that Batwa culture and lifestyle (in addition to poverty) were the key causes of poor learning among the Batwa, community leaders and parents noted rampant discrimination against the Batwa people, government rejection and social exclusion.

Multicause factor diagrams provided rich and nuanced depictions of the prevalent perceptions among key stakeholder groups and contributed to the overall study findings and recommendations. Most importantly, the workshops and resulting maps helped identify key leverage points that would allow USAID to develop effective interventions (IBTCI, 2020,

p. 6). The study produced separate programming recommendations for each of the three study population groups.

LESSONS WE LEARNED

This pilot taught us a lot. In this section, we summarise some of the questions we have asked ourselves over the two years of this case study.

Did we Pilot a System Diagnostic?

The pilot diverged substantially from the plan. We had originally planned to pilot two countries together and test two different system diagnostics. We had built a three-step process, but only step two was completed in Rwanda and the DRC. Data analysis intended to inform the in-country conversation was not used. We had hoped for a holistic overview of education outcomes, but even taking the particular regions of a country, this proved too abstract a level of analysis, so in both cases, more specific questions were developed – questions that moved us away from the 'system diagnostic' goal. We found that it was necessary to set boundaries around a subsystem or issue of interest within the system, and even this narrower view revealed significant complexity.

Did systems thinking help us get to a deeper understanding than we had found possible before? Did this reveal anything to the participants that they didn't know already? To this, we give a resounding yes. In both countries, the participants – and funders – learned a lot of new and unexpected information and perspectives on education challenges. Our evaluation in Rwanda reported that the participants found the workshops to be effective, and they were unanimous that the workshops increased their understanding of education issues in Rwanda. No independent evaluation was conducted in the DRC, but feedback to USAID was anecdotally similar. The mapping exercise in Rwanda helped stakeholders see education issues holistically rather than considering the influencing factors in isolation. The participants also appreciated the opportunity to hear other peoples' views and experiences related to the Rwanda education system. The follow-up evaluation visit in February 2020 confirmed the profound effect the workshops had on the participants. Even months after the workshop, the participants were able to clearly demonstrate how they could use the learning from the workshop in their everyday lives. However, the findings from the system diagnostics exercise were not shared with any of the participants, and they were not aware of how their

findings were used because the findings informed procurement design, thus becoming embargoed. This is something to reflect on for future attempts, but a key insight from systems thinking that differs from more rational, linear thinking is that it is the process for participants where the value lies, not in the sharing of the findings or mappings.

Was the System Diagnostic Process Palpably Different From a More Traditional Sector Assessment?

Although both diagnostic studies that have been described in this chapter used the same methodological approach, the overall scope and scale of the diagnostics in the two studies were different. The pilot in Rwanda was limited in time and resources and implemented influence mapping workshops in conjunction with system diagramming. The system diagnostic in the DRC, by contrast, was part of a larger needs assessment study employing multiple data collection methods. The DRC study team triangulated the results of the influence mapping workshops with data from KIIs and FGDs, identifying critical information derived from different methods, contributing to an improved understanding of the unique advantages of these methods and bringing about recommendations for ideal sequencing.

Despite these differences, both system diagnostic studies were found to be an effective alternative to traditional assessment methods. Both studies produced insights that otherwise may not have emerged as clearly or prominently through other methods. Both studies were significantly cheaper and faster to implement compared with the more traditional methods. Finally, both studies informed not just the commissioner of the study and the research team, but also the participants. The participatory nature of implementing systems diagnostics allowed for a much more involved, transparent and collaborative process of meaning-making than more traditional methods, such as KIIs or FGDs, ultimately leading to increased stakeholder ownership of the findings and resulting actions.

The DRC system diagnostic was particularly informative in comparing the application of systems thinking methods with more traditional qualitative data collection methods because both approaches were used. Table 10.1 was developed by the MECC team for the presentation to USAID as a way to illustrate the advantages of different methods, with the text set in bold type illustrating an advantage.

As Table 10.1 illustrates, systems thinking approach-based diagnostics can be advantageous over the more traditional methods. However, their

Table 10.1   *Comparison of traditional methods with systems thinking approaches*

Key informant interviews/focus group discussions	Systems diagnostic workshops
Resource intensive	**+ Fewer resources**
Lengthy process	**+ Quickly implemented**
Skilled interviewers and note-takers	Skilled facilitator + note-takers
Preparation and logistics: identify participants, multiple sites, travel	Preparation and logistics: identify participants, workshop site, travel
+ Broad focus	Narrow(er) focus
Multiple tools	**+ Fewer tools**
More time needed for analysis and reporting	**+ Less time needed for analysis and reporting**
+ Student perspective included	No students included
	+ Discussions provided new findings

implementation is not without their own challenges that can significantly impact the validity and reliability of the results. First, participant selection is of particular importance. Recruiting participants who bring a diversity of perspectives can be a challenge, and in some circumstances it may be difficult to ensure an adequate representation of all key stakeholders. In particular, neither the Rwanda nor DRC diagnostic studies involved students. Engaging minors in systems diagnostics workshops would need to be explored from an ethics perspective, and clear guidelines need to be developed, but student voices should be heard wherever possible.

Second, the number and duration of workshops should reflect the complexity of the issue being interrogated. Unlike the more traditional methods, more complex issues that involve more diverse stakeholders and inspire differences of opinions will necessitate more extensive data collection. The time limit has a great impact on how many layers of causes and factors the participants have time to explore and how much time the workshops can allow for integrating these competing perspectives.

Finally, system thinking practitioners tend to combine the use of different mapping tools to enhance the validity and reliability of the findings. Additionally, the DRC diagnostic study suggests that some combination of systems thinking methods and more traditional methods can be advantageous in some situations (as we had anticipated at the start of the pilot, with the design of a more traditional 'step one' before the systems thinking 'step two'). Further exploration of how to best match diagnostic

questions with the most suitable tools from the systems thinking toolkit is needed.

Despite these caveats, the two diagnostic studies presented here have demonstrated the utility and appropriateness of systems thinking tools for the successful implementation of diagnostics in the international education sector. Both studies produced robust sets of recommendations that were found to be actionable by stakeholders and were subsequently integrated by USAID Missions in project design.

Was this Too Direct an Effort to Translate Global Dialogue into National Practice?

We found our positioning at the headquarters of global aid agencies to be the wrong point from which to initiate this effort. Our evaluation found that USAID Rwanda was closely engaged and, thus, benefited from the pilot, but this was true to a lesser extent for DFID. The evaluation concluded that 'relevant ministries in Rwanda did not own the system diagnostics agenda or process' (Karki et al., 2020).

Did this Way of Doing it Bring Others Along, Building 'Coherence' of Understanding of the Problem and What Needed to Change?

Yes, more so than consultants flying in to write a report, but we were very much aware that we and our contractors still represented global aid agencies.

Did We See Anything Change as a Result of These System Diagnostics?

We were not so naive as to expect behaviour change from this exercise, which ultimately involved attending one or two workshops for a small number of people. However, we had hoped to see some mindset shifts or even a decision or two that was different than we had expected. Behaviour change takes time, but with these two pilots, we planted a seed. The USAID staff members who participated are converts and use systems thinking tools in their work. USAID has integrated systems thinking into a new award mechanism for primary research and released internal guidance for USAID staff on how to do this. This is already changing the behaviour of USAID staff (and beyond), and USAID is

optimistic that this diagnostic approach will have a transformative impact on the evidence ecosystems they engage with.

We differ slightly in our final conclusions, and our concluding comment is that we could apply some of these systems thinking reflection tools to our own systems thinking exercise.

For Laura, this was far from the pilot that she had anticipated at the outset and was much, much harder to do than anticipated. It was harder to convince our colleagues that it was needed: to show them that better problem identification was needed, that a coherence of the response was needed and that stronger accountability mechanisms were needed. It was harder to navigate the complex politics of aid from our headquarters' standpoints to identify true government demand for this from where we sat. It was hard to bring not only untested tools into a high-stakes education sector planning conversation, but a framework around these tools that was explicit – and enthusiastic! – about complexity, messiness and flux. Ultimately, the tool was not piloted as envisaged. Systems thinking might appeal to us in explaining that there is no such thing as a solution, but this does not work for those who so desperately want there to be one.

Elena, however, has a much more positive set of reflections, perhaps because she has gone on to apply these tools in partnership with several other missions and has seen widespread take-up in USAID. She reflects that the hardest part was finding the right opportunity to test this approach, but that in practice – and Elena participated personally in the workshops, whereas Laura did not – the systems thinking tools were easy to apply. The tools prompted us to rethink how we engage with other actors and how to break down sectoral boundaries. The positives are clear and implementing a system diagnostic is doable. However, the journey does not stop at completing the systems thinking exercises; systems thinking is also needed to translate the messy and cross-sectoral findings into practice and programming.

REFERENCES

Andrews, M., Pritchett, L., Samji, S. & Woolcock, M. (2015). *Building capability by delivering results: Putting Problem-Driven Iterative Adaptation (PDIA) principles into practice.* OECD. https://www.oecd.org/dac/accountable -effective-institutions/Governance%20Notebook%202.3%20Andrews%20et %20al.pdf

Department for International Development (DFID). (2018). *DFID education policy: Get children learning.* https://www.gov.uk/government/publications/ dfid-education-policy-2018-get-children-learning

Faul, M. V. (2016). *Scoping study: Education systems diagnostics tools*. DFID.

Global Partnership for Education (GPE). (2019). *GPE results framework for 2016–2020*. https://www.globalpartnership.org/content/gpe-results -framework-2016-2020

Harris, D., Mcloughlin, C. & Wild, L. (2013). *The technical is political – Why understanding the political implications of technical characteristics can help improve service delivery*. Overseas Development Institute.

International Business & Technical Consultants, Inc. (IBTCI). (2020). *Literacy Landscape Assessment in the Democratic Republic of Congo: Assessment Report*. USAID. https://www.edu-links.org/sites/default/files/media/file/ PA00X79F.pdf

Karki, S., Pontalti, K. & Majeed, Z. (2020). System diagnostics in education. Unpublished report.

Laterite. (2018). *Understanding dropout and repetition in Rwanda.* http:// www.rencp.org/wp-content/uploads/2018/09/DROPOUT-STUDY-FULL -REPORT.pdf

Strategic Program for Analyzing Complexity and Evaluating Systems (SPACES). (2019). *Rwandan primary students learning to read Kinyarwanda: A systems diagramming approach*. USAID. https://linclocal.org/wp-content/uploads/ 2020/04/Rwanda_Education-System-SPACES-FINAL-REPORT.pdf

Teskey, G. (2017). *Thinking and working politically – Are we seeing the emergence of a second orthodoxy? Abt Associates.* https://www.abtassociates.com/ insights/publications/white-paper/thinking-and-working-politically-are-we -seeing-the-emergence-of

USAID. (2016). The 5Rs framework in the program cycle. https://usaidlearninglab .org/sites/default/files/resource/files/5rs_techncial_note_ver_2_1_final.pdf

Wild, L., Andrews, M., Booth, D., Valters, C. & Dempster, H. (2016, 16 December). *Doing development differently: Who we are, what we're doing and what we're learning*. ODI. https://odi.org/en/publications/doing-development -differently-who-we-are-what-were-doing-and-what-were-learning/

World Bank. (2018). *Schooling for learning – Strengthening resilience of education in Rwanda*. https://openknowledge.worldbank.org/handle/10986/30978

World Bank. (2019). Global education policy dashboard. https://www.worldbank .org/en/topic/education/brief/global-education-policy-dashboard

PART IV

Conclusion

11. Conclusion: beyond silver bullet solutions

Moira V. Faul and Laura Savage

If you started this book believing that systems thinking is the singular solution to support quality education for all, then we hope that you are now questioning whether one 'solution' is really possible and are maybe even feeling okay with that. We also hope you have gained a sense of how to approach complex problems within education systems, the roles of experts and participants, and the roles we all play as actors within various interdependent subsystems (development, humanitarian, funding, policy, practice and research, among many others).

We set out to curate this volume after having worked together on the theory and practice of systems thinking in education for over a decade, with the hunch that there was a wave of exciting applications of systems thinking happening in international education. The chapters in this book showcase empirical applications and diverse theories of systems thinking, highlighting a variety of ways in which to think about education as a system, at a range of various scales and from a variety of perspectives. The contributors to this volume describe various tools – academic and practical – that we can consider using in our work, whether as an evaluator, funder, policymaker, practitioner or researcher. Yet we also found that the field remains nascent and constrained.

Even though they adopted different theories or conceptual frameworks, all the authors started from the premise that applying systems approaches to education can hold significant implications for research, policy and practice and improving learning for all. This conclusion offers a synthesis across the wealth of theories employed and empirical data and analyses offered in this volume's chapters. We draw on these to offer reflections on the implications of systems thinking for what we might do – in whatever role we play in the effort to achieve learning for all.

REFLECTIONS ON SYSTEMS THINKING IN INTERNATIONAL EDUCATION

Systems Thinking – or Aspects of It – is Growing in Appeal

Over the past decade, the word system has become widespread in global education discourse. There is openness to a lot of the core ideas: that education processes, policy and decision making are complex and dynamic; that relationships are important; that there are unintended consequences to what we do that flow into feedback loops that can hold systems in positive or negative patterns; and that learning and equity are two of the many emergent properties of education systems. There is widespread acceptance of the disconnect between planning and practice (such as GPE, 2021), based on research and lived experiences which show that the best intentions and even the most faithfully implemented projects do not necessarily translate into the hoped-for outcomes. The chorus of voices asking 'Why not?' and 'We need a revolution' grows louder (Goddard et al., 2022). School closures in response to the COVID-19 pandemic have generated a truly global conversation about student learning loss and educational inequality. The commitment to 'build back better' implies a willingness to work differently and take on radical ideas (UNESCO, n.d.).

The chapters in this volume showcase this shift. Almost all the chapters provide analyses of case studies in which researchers, policymakers and practitioners have willingly – even enthusiastically – used systems thinking tools in the design, implementation or evaluation of an activity, whether in classrooms and schools or national priorities and international aid. Thus, this collection represents 'a testimony to the feasibility and advantages of thinking systemically about educational interventions' (Parra and Edwards, Chapter 9, p. 184).

Although the Word 'System' is Increasingly Used, Systems Thinking Is Not

Despite the enthusiasm with which major actors in international education use the word 'system', we have not found evidence of a widespread uptake of systems thinking – the mindset through which a complex education problem is understood as a system – much less the practices that follow from that understanding. Instead of systems thinking and prac-

tices, linear thinking and practice abound. Education sector plans, logical frameworks demanded by funders and school improvement plans around the world continue to be riddled with assumptions of linear trajectories from receiving funds to implementation to outputs and finally to outcomes. Countless strategy documents aim to achieve 'systems change' by testing 'what works' and then applying this 'at scale'. Rather than taking systems thinking seriously, many 'build back better' advocates propound magical ideas that promise to immediately translate into fast and mass success. Yet this has never worked before; why would it work this time? Thus, incentives against adopting systems approaches are more common than those that encourage their adoption.

Undertaking a systems thinking process tends to be experienced as useful and positive. For some, however, the resulting picture of complexity can be dissatisfying, particularly to those with the drive to do and to change. However, doing something absent a systems approach may be counter-productive – or at best may have no effect at all. Using systems thinking to work through the very real messy complexity of education systems allows for the identification of a potentially useful way forward – for now, at least – that can evolve the system in a useful direction.

Systems are approached throughout this volume from a wide range of theoretical approaches. As the editors of this volume, we made a deliberate choice to showcase rather than sidestep this complexity. Several frameworks are presented through which to holistically understand education systems and the process of system change, bringing in core concepts from complexity theory, inclusive systems development theory, sociological systems theory, economics and political science, explaining behaviour and consequence through the analysis of relationships between different sets of education system actors. Simultaneously, for several authors, this is their first time applying an explicit systems lens to their work. For example, theory emerging from deep research on country case studies has established the value of understanding that children can exist in one or more different zones of exclusion. These zones come to life under a systems lens, illuminating the dynamic nature of a child's journey through – in and out of – school and encouraging a policy focus on transitions as moments that can be anticipated or diverted. The authors who report on their efforts to improve the intrinsic motivation of teachers by connecting them in professional peer networks did not start from a systems approach. Instead they found that systems thinking adds a valuable perspective to their analyses and future plans for the project. Additionally, the collective of funders, practitioners and researchers

involved in a continent-wide project to implement ability-based peda-gogy reflected that 'while very tempting to package into a series of "best practices" that could be replicated elsewhere, we believe that the primary lesson our case study illustrates is that successful work with systems is guided by a series of principles, while the choice of actions and processes is thoroughly adaptive over time' (Lipovsek et al., Chapter 2, p. 42). Thus, not all authors sought to explicitly prove or implement systems thinking theory; rather, they found value and sense in systems thinking concepts as they reflected on their findings or practice.

And Systems Practices Are Used Even Less Than Systems Thinking

Our hypothesis is that the problem of achieving learning for all is not com-plicated but complex. This demands a shift in practice. Systems thinking enables us to see education systems as they are: complex, dynamic and not amenable to one-size-fits-all technical solutions. Systems approaches recommend iteration and treating short-term failure as data to learn from, in order to adjust and fare better in the next iteration (of many more to follow). Many system actors whose hands are on the levers of systems change interpret cycles of learning and improving as failure and risk, whether governments (in high- and low-income countries) who want to see policies succeed in their short-term political cycles; funders who are accountable for public or private money; or practitioners and researchers who depend on external funders and their demands for positive results in the short term. In this volume, we invited researchers and practitioners to write about what they do, show how they grapple with layers of com-plexity, and write down the how and why of systems thinking. Most prac-titioners lacked the time or confidence to write a 6,000-word academic piece, and some who could feared the consequences for their funding or reputation if they wrote about what could be construed as failure, even if they reported learning and improvements resulting from that. Systems practice is the behaviour that emerges from systems thinking and that indicates more appropriate ways of taking action, actions that are more likely to move us closer to our system goals.

SO WHAT DOES SYSTEMS THINKING MEAN WE DO DIFFERENTLY?

Systems thinking offers a framework through which to approach the two questions set out in the introduction: why technical solutions might not

work and how to act differently. There are some useful behaviours that we can aim to adopt.

Define the System You are Working With

The majority of this volume's authors take a broad definition of systems as (1) elements: human, social and technical, (2) the relations between them that structure their interactions, and (3) existing within specific contexts that affect and are affected by those elements and their relations. Whether you are working with or researching a handful of schools, a network of teachers or a whole country's education sector, you need to be clear about where you are drawing the system's boundaries. In mapping the TVET system in Cambodia, for example, the system boundaries were drawn around all the key organisations and sectors identified as playing a role in perpetuating the problem while remaining agnostic as to whether these organisations were considered local, district or national. This allowed for the identification of where best to intervene in that system to improve the system outcomes for the affected youth, including changing the offer made by the provincial training centres and working towards more productive relationships between the national government and provincial training centres and other partners. Analysis that would have insisted on either methodological nationalism or focused only on one level would have missed both the elements at the other levels and the complex entanglements between them that were producing system outcomes that proved to be counter-productive for the youth the system was intended to serve.

Part of the reason for setting boundaries around a system is to be able to identify the external dynamics and influences that shape and change it. Across the different theories, approaches, empirical contexts (schools, districts, countries) and levels of analysis collated in this book, all authors have conceptualised education systems as open, not closed, to outside influences. This is an important move away from a mechanistic understanding of education systems as closed 'plumbing' models, allowing the researcher or decision maker to identify the other systems that may impinge upon an education system (for example, environmental or economic or families or donors) and the demands they may make on the education system. One structural analysis offered – of education as a functional system that is cognitively open (if operatively shut) to other functional systems (such as politics or law) – shows how, when and why reforms may travel from one jurisdiction to another. Even classrooms

are open to wider social systems: the specific types of inequalities and marginalisations that may be found outside of schools can influence the actions and interactions inside that classroom, between teachers and pupils, and among pupils. Different contexts are imbued with different social norms and values, historical marginalisations and economic inequalities, which any system intervention needs to consider.

Another contribution of systems approaches lies in their application to a range of different scales or levels. The micro level can be conceptualised as those elements inside the system, the relations between them can be seen as forming the meso level, and the organisational settings in which they work can be seen as the macro level, all of which are embedded in broader social, economic and environmental contexts. Improvements in local classroom practices (micro level) were brought about by seeding and supporting communities of practice (meso level) that work in different schools (macro level) located either in the same city (Delhi), different regions (five regions in Kenya) or nationally (Rwanda) that were funded by a consortium of donor agencies. Thus, systems approaches can be applied at and to different scales and levels: from classrooms and professional networks to district, national and global decision making. Setting your conceptual boundaries around a system allows you to focus your analysis and practice in the right place; it also avoids the otherwise panic-inducing feeling that the world is one giant incomprehensible system that you cannot begin to grasp.

Support Collaborative Problem Identification

The next step is to better understand how the system works. Within this system, what are the processes, structures and incentives that hold the system where it is? Who are the actors with the most power, and what do they want? Where are the potential leverage points that could move the system towards better supporting quality learning for all? Starting with understanding the problem and why previous reform attempts have affected change is better than starting with a solution and assuming that it is going to work. This step can be done externally (as donor-funded analyses often are), but it is far more effective when done collaboratively so that the range of actors within an education system gain perspective together on the diversity of goals and objectives that each other holds, on the different ways they understand change to happen and on the competing views on what to do about it.

For example, working with the people within the system was critical to the success of the Teaching at the Right Level initiative in Zambia and India. The global system actors aligned funding and support behind the priorities of national ministries on piloting an innovative approach to improving literacy through a South–South collaboration. In the case of the assessment of a large-scale intervention in Colombia, the authors deliberately sought out the perspectives of the programme designers and implementers to understand the system they were working with and increase the acceptance of their study among those practitioners whose actions were seeking to improve the system under investigation. The experience of testing a system diagnostic in Rwanda and the Democratic Republic of Congo showed that using systems thinking tools (such as causal loop mapping) yields new insights, which changed funding decisions and programme direction. Reflections on that pilot make clear that this process of collaborative problem identification is not a straight-forward or efficient process. It is effective, however: instrumentally, endogenous change processes are more likely to be sustained and sustainable; normatively, the demand for representation and voice should not be denied.

Be Conscious of Your Role: Within, Outside Of and as Part Of That System

Systems approaches recognise that systems permit 'equifinality (different pathways to the same goals) and multifinality (different goals and overlapping pathways)' (Lewin, Chapter 6, p. 132). This means that systems analysts and practitioners must accept the limitations of their ability to drive systems change in one direction or down one particular path. Instead, incremental and evolutionary change is possible and desirable by accepting (rather than ignoring) the flux and complexity inherent in systems change. In other words, no single actor or group of actors can change a system. We can identify the levers that might nudge change in the direction we want and then examine the feedback loops and unintended consequences to work out where next to engage.

There was strong recognition of the unpredictability of system change across the chapters but limited self-reflection on author positioning or perspective within a system. Only when a specific question was put to global education aid actors on their role in supporting system change did reflections on this point emerge. Here, there is thoughtful and candid contemplation of the difficult balance between recognising that change

needs to come from within a system and knowing as an 'outside' actor that you might be able to help shift some of the obstacles impeding that change. Several authors recognised the importance of internally driven change and work to achieve this local ownership. For example, researchers, funders and practitioners worked together with the Government of Zambia to co-create an approach to teaching at the right level (TaRL), adapting lessons from experience in India. An interesting challenge to this approach comes from the multicountry RISE programme. The RISE framework deliberately does not present actors within a system as 'insiders' or 'outsiders', but instead emphasises that there must be coherence across all actors towards learning for incentives in the system to change, and in their relationships since they affect system dynamics and, therefore, change. In addition to reflecting on your role in one system, it is also critical to interrogate how the dynamics of other systems you are part of (government/civil society or donor/recipient) may influence your behaviour and interactions inside an education system.

Design Policies and Programmes Focused on People and Dynamics Rather than Interventions and Numbers

The system elements studied in these chapters include people (teachers, mentors and headteachers), and also provincial training centres, central government and private sector firms; schools, regional education secretariats and the national ministry of education; a group of donors aligned with a national government priority on evidence-based teaching reform working together on a pilot project with teachers in schools; a broad range of different types of elements, including human, conceptual, social, political, environmental, economic; citizens, education workers at the chalkface; and educational and other authorities, to mention just a few. In systems approaches, these elements are not treated as if they exist in isolation because they do not. Multiple types of relationships between diverse groups of stakeholders matter.

Relationships between these elements shape education system dynamics. The relations studied in this volume's chapters are characterised by accountability, but also by trust, challenge, dialogue, collaboration or shared purpose. These relations may be horizontal (between peers) or vertical (within a hierarchy). The connections between system elements can be correlation based or causal, identified through evidence from academic research and/or by the multiple perspectives of system actors. Beyond educational settings, the web of relationships in which each

learner is embedded is also crucial to their navigation of the schooling system in ways that support learning. In many chapters, the systems intervention examined was designed specifically to influence system relationships, whether to strengthen them or make them more positive or productive. In many chapters, through improving these relations, professional practices and cultures improve and become better aligned – or coherent – to achieving the desired system outcome.

Do Not Assume that Scale Means Big

Inherent to systems thinking are the concepts of context and constant flux. Even when working towards goals that are shared – when there is strong collective problem identification and incentives designed towards achieving a particular objective – the system may evolve in ways that are unpredictable. Achieving system evolution towards shared objectives requires alignment of the intervention to the goal and to the system into which it intervenes. The authors in this volume have unanimously presented this view. Applying Luhmann's sociological systems theory shows that reforms cannot be simply copied from one system to another; certain system factors must align for policy innovations to travel. Exclusion traps for children who are marginalised in different zones of transition require different types of remediation. Vocational education project design in Cambodia was carefully aligned with the dynamics of that system, including buy-in for policy changes. If an intervention is rolled out at scale it may be incoherent with the incentives within the systems into which it is scaled. Recognising this, systems approaches move beyond designing a single intervention that is implemented – or scaled into all contexts. Instead of considering the fidelity of scaling, attempts to repeat an intervention in a new place must be considered as trialling, with tight feedback loops of test–learn–adapt; interventions must be adjusted to respect the local context and distinctive response to the intervention that arises as it is implemented in that context. Even better, the goal should also be to find local innovations and interventions that have been shown to work and have local buy-in.

   Yet none of the authors here challenges the very notion of 'large scale'. Perhaps this is the case because the learning and equity crisis is so vast and severe that a massive response is felt to be needed. However, the aspiration towards scale can diminish the quality of education being provided if it is planned poorly and driven from the outside. Systems thinking encourages thinking about scale as appropriate to the context:

an intervention may be appropriately scaled when it has worked for just a few schools that share a particular dynamic.

Enable Test–Learn–Adapt Approaches that Work with Tight Feedback Loops

Working towards change that can enable learning for all will require being 'comfortable with uncertainty and being able to change course' (Lipovsek et al., Chapter 2, p. 41), which places value on sustained change. Rather than holding onto the aspirations of what the system or a particular 'solution' should be, this volume shows the importance of starting where the system is and working with the system as it is. Doing the same thing in different places, at different times or with different people does not yield the same result. In contrast to conventional linear interventions and 'rational' theories of development, the results of any systems intervention are recognised to be – and therefore treated as – inherently and essentially unpredictable and nonlinear. Any intervention into a complex system (such as education) needs to be attentive to unexpected outcomes, changes in relations inside the system, emergent outcomes, and the interactions with its context or other systems. For example, in implementing the TaRL approach in Zambia, the teachers were given autonomy over grouping children at different levels, with several iterations over time in which the children would move on or stay in the same group, as relevant to their learning needs. A shift in perspective is needed: from short-term planning goals to longer-term system evolution; from desires for neat and efficient one-size-fits-all solutions to evolutionary, incremental effectiveness in the real world. Hence, it is critical to maintain a tight feedback loop between intervention action, data on that action and responses to those data to keep the system evolving towards the goal. The role of data is not to rank, report or punish but rather for system actors close to the intervention to respond to these data in ways that keep the desired evolution on course. Thus, sustained change in the desired direction is achieved incrementally and follows a nonlinear path that may be less direct, but a path that is more likely to move the system towards the desired objectives. The successful system interventions in these chapters were guided by a set of principles of working with systems, using a targeted strategy with respect to differences in each context and evolving that strategy as the system into which it was intervening also evolved.

Adopt a Glossary of Terms that You Use Consciously

We hope this volume has prompted you to think about how you use the 's' word, and also to reflect on other words often associated with system change. 'Scale', for example, does not mean big; a project can be at scale when it has reached the number of schools, teachers and students for whom it works. 'Solution' is another word we would urge you to use with caution; it prompts a mindset of simple, linear pathways that run from funding to implementation to success. There is no single solution to the global learning crisis, and the sooner we appreciate this fact and adjust expectations, the sooner we stop looking for quick fixes. The authors here have applied different theoretical lenses and have not all adopted the same glossary. Yet each unpacks at least some important terminology, taking us one step closer to developing a common language to describe how systems evolve and how to work with them.

WHERE COULD EDUCATION SYSTEMS RESEARCH GO NEXT?

More empirical cases are needed for learning from existing practice in international education and to continuously refine systems theory or hypotheses. Researchers could also bring learning from other sectors, such as studies into systems approaches as they are used and researched in agriculture, energy or health (Borman et al., 2022; Brouwer et al., 2020; Ingram, 2011). We find more fluency in systems thinking and practice in other sectors more related to the natural world (agriculture, climate and ecology), from which education can adapt valuable lessons. In addition, methodological refinement and training is required (particularly in contrast to the methodological nationalism so prevalent in education research) to build researchers' competence and confidence in applying system methods in their work or in analysing systems approaches undertaken by others. For these advances in research to occur, funders and academic institutions need to build more flexibility in their institutions and processes to support systems thinking and systems thinkers.

Arguably, the Global North is rediscovering thinking in systems, but this is hard-wired into many Indigenous knowledge systems (Kimmerer, 2013; Shiva, 2013). The historical hegemony of knowledge from the Global North over other forms of knowledge was 'achieved at the cost of tremendous silencing, parochial legitimation procedures and, most of all, the deterioration in social status for most of humanity, including women

and non-Western cultures' (Odora Hoppers, 2002, p. 27). Now that systems approaches are becoming more mainstream in the Global North, it would be a category error to attempt to frame these as the latest in Western thought, rather than acknowledging the long histories of systems approaches in other knowledge systems and practices.

FINAL REFLECTION

This volume provides a wealth of cases, data and analysis on how system change may be pursued from diverse positions inside and outside education systems and in ways that may contribute to improvements in access, learning and equity. The chapters provide several examples of efforts to apply systems thinking to improve education quality at several different scales (classroom, school, district, national and global), as well as within the same level (such as teacher communities leading from the middle or donors leading a shift towards putting system change front and centre). Across the chapters, we see how collaboratively identified factors internal and external to education affect processes and outcomes. We find fertile ground for further debate on how to support system change, as well as pointers for future research agendas.

Although this is a book about systems thinking in education, it is not a manifesto. We find the language of 'systems thinking' useful. However, we are keenly aware that many different languages may translate into similar practices. We are not promoting systems thinking as a silver bullet solution that everyone needs to get on board with. There are communities in international education which are more comfortable using the language of 'agile' approaches, or 'human-centred design' (for example, UNICEF, 2019), or of 'doing development differently' (Wild et al., 2016), of 'people-first development' (People First Dev., 2021) or of 'thinking and working politically' (TWP, n.d.). Despite using different vocabularies, we see crossovers and synergies in practice. It would be unhelpful to set ourselves up as distinct and adversarial brands of change; indeed, it would be a betrayal of the principles of systems thinking to do so.

The world is complex, not complicated, and so are education systems and processes. Acting as if they are not merely creates more unintended consequences without moving us any closer to achieving equitable quality learning for all. This volume seeks to continue – while contributing to – the impetus to move beyond simply using the word 'system' instead of sector, to taking seriously the promise and practice of systems

practices to improve – and improve research into – the functioning of education systems at every level and scale to deliver on the promise of learning for all.

REFERENCES

Borman, G. D., de Boef, W. S., Dirks, F., Gonzalez, Y. S., Subedi, A., Thijssen, M. H. ... & van Berkum, S. (2022). Putting food systems thinking into practice: Integrating agricultural sectors into a multi-level analytical framework. *Global Food Security*, *32*, 100591.

Brouwer, I. D., McDermott, J. & Ruben, R. (2020). Food systems everywhere: Improving relevance in practice. *Global Food Security*, *26*, 100398.

Goddard, C., Chung, C. K., Keiffenheim, E. & Temperley, J. (2022). *A new education story. Three drivers to transform education systems.* Big Change. https://big-change.org/new-education-story/

GPE. (2021). *Draft guide for enabling factors analysis for GPE system transformation grants.* https://www.globalpartnership.org/content/draft-guide-enabling-factors-analysis-gpe-system-transformation-grants

Ingram, J. (2011). A food systems approach to researching food security and its interactions with global environmental change. *Food security*, *3*(4), 417–431.

Kimmerer, R. W. (2013). *Braiding sweetgrass: Indigenous wisdom, scientific knowledge and the teachings of plants*. Milkweed Editions.

Odora Hoppers, C. A. (Ed.). (2002). *Indigenous knowledge and the integration of knowledge systems: Towards a philosophy of articulation*. New Africa Books.

People First Dev. (2021). Collective leadership as a path for sustainable development. https://peoplefirstdev.org

Shiva, V. (2013). *Making peace with earth.* Pluto Press.

TWP. (n.d.). Thinking and working politically. Community of practice. https://twpcommunity.org

UNESCO. (n.d.). *Futures of Education*. https://en.unesco.org/futuresofeducation/

UNICEF. (2019). *Human-centred design.* https://www.unicef.org/innovation/media/5456/file

Wild, L., Andrews, M., Booth, D., Valters, C. & Dempster, H. (2016, 16 December). *Doing development differently: Who we are, what we're doing and what we're learning.* ODI. https://odi.org/en/publications/doing-development-differently-who-we-are-what-were-doing-and-what-were-learning/

Index