Online Appendix

Review of Systems Approaches in Health, Education, and Infrastructure

Zahra Mansoor and Martin J. Williams

August 2022

1. Introduction

In recent years, the idea of taking a systems approach to understanding public service delivery has gained currency with academics and policymakers alike (OECD, 2017; Bandiera et al., 2019; Besley et al., 2021). This online appendix reviews the state of systems approaches in three sectors where such approaches have been increasingly employed to understand service delivery: health, education, and infrastructure. Systems approaches go as far back as the 1930s (Jackson, 2009), with some of its initial ideas captured by the works of Von Bertalanffy and his General Systems Theory (1972), that was mainly based on the sectors of biology and ecology. This evolved into other sectors such as engineering through cybernetics and systems engineering in the 1940s (Hall, 1962), and later into more quantitative computer modelling approaches leading to the field of systems dynamics (Forrester, 1961; 1968). Around the 1970s, debates around how to capture social elements of the world into systems approaches led to the emergence of *soft systems* approaches (Checkland, 1999), which naturally saw the use of systems approaches in more human-centered sectors such as health and education.

Whereas the main text of the paper takes a synthetic and integrative approach across sectors, in this online appendix we provide a sector-by-sector view on the state of systems approaches in our three focus sectors of health, education, and infrastructure. This appendix will thus be of interest to readers interested in their particular sector or in how the shape of systems approaches differ in other sectors. It also provides readers with a fuller picture of the literature review on which the conceptual argument in the main text of the paper is based. In compiling this review, we used a combination of citation-tracing from foundational texts and review papers (such as e.g., Gilson, 2012; Carey et al., 2015; and Hanson 2015 for health; Pritchett, 2015 for education; Saidi et al., 2018 for infrastructure), keyword searches in online databases, and suggestions from sectoral experts. The result is not a systematic review in the formal sense of the term, but nevertheless provides a detailed and consistent picture of the state of the literature in each sector.

Given the diversity and range of systems approaches in these sectors, we organize our review by two types of systems approaches within each sector: macro-systems approaches and micro-systems approaches. The former are primarily concerned with looking at the entire system as a whole, and the use of systems approaches to understand the collective coherence of a set of policy interventions with each other as well as various other elements of context. The latter focus in on a single policy intervention, with emphasis on the use of systems approaches to understand not only whether the policy in question works, but also how it interacts with other elements of the system.

Section 2 reviews the state of the systems literature in the health sector, followed by the education and infrastructure sectors in Section 3 and Section 4.

2. Health Sector

2.1 Motivation, Definition, and Scope

Health systems research developed as a field over more than decade ago as a way to understand complexities, interrelationships, and structural constraints within health systems. Limited success of emergency responses to major health epidemics around the world (such as the West African Ebola outbreak in 2012-2014) underscored the need for coordinated action across various system actors such as policymakers, health service providers, health recipients, whilst taking various contextual realities into account. In addition, the sluggish progress of key health indicators around the world despite significant investments in a range of narrow interventions brought attention towards structural weaknesses in health systems (Travis, 2004; WHO, 2007).

These factors highlighted the urgency for research focused on health systems.

The World Health Organization (2007, p. 2) defines a health system as consisting of "all organizations, people and actions whose primary intent is to promote, restore or maintain health." The Alliance for Health Policy and System Research (2011), a health systems research forum initiated by the World Health Organization (WHO) in collaboration with the Global Forum for Health, defines *health systems research* (HSR) as the production of knowledge that is geared towards understanding how societies organize themselves to achieve health goals.[1] This definition implies a focus on how the 'whole' system functions instead of a narrow focus on any single aspect of its individual components (Hanson, 2015).[2]

Three key features summarize the scope of systems research in this field. First, health systems research has a specific focus on real world issues. It aims to address questions which are practically faced by countries within the health sector. Second, it is multidisciplinary drawing on disciplines such as economics, sociology, anthropology, political science, public health, and epidemiology. This is closely linked to the first point as real-world issues about health systems could be of the 'why', 'what', and/or 'how' nature. Hence the disciplinary or methodological grounding of research is determined by the question of interest. Third, the research is applied with a unique focus on policy with the goal to influence policy. This implies that research with respect to how policy is made and implemented is a key area of research focus for the field (Mills, 2012; Gilson, 2012).

---

[1] The Alliance for Health Policy and System Research (AHPSR), comprising of health practitioners and academics from around the world, has taken instrumental steps in defining the scope, boundaries and agenda of the field through several publications.

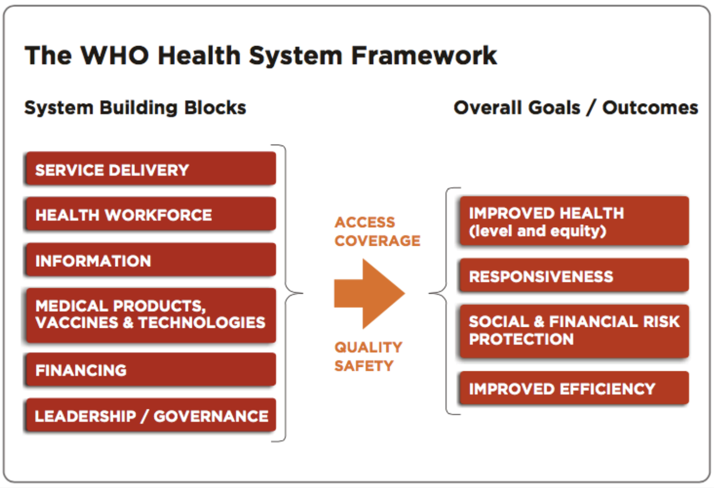
[2] The Alliance included the word 'policy' in what is commonly known as health systems research and renamed it to 'Health Policy and Systems Research (HPSR)' to capture two key facets of the field which do not clearly come through in the definition – first to highlight the importance of social and political realities within a health system and second, to recognize the applied, policy, and question-driven nature of the field.

2.2 Macro-systems Approaches

Macro-systems approaches in HSR focus on "whole" systems that aim to understand coherence, coordination, and interconnection of various policies within systems. These approaches vary in the level of specificity with which they define relationships between health system components. In this sub-section, we illustrate this variation through a range of different examples.

One type of macro-systems approaches merely describe health system components in different ways, such as in terms of its functions, stakeholders involved, or hierarchical levels. The seminal WHO framework defines the health system as comprising six key functional building blocks - service delivery, health workforce, information, medical products (including both vaccines and technologies), financing, and leadership and governance – and links them to the broader health system goals (WHO, 2007).

Figure 1: WHO Health System Framework



Source: De Savingy and Adam (2009)

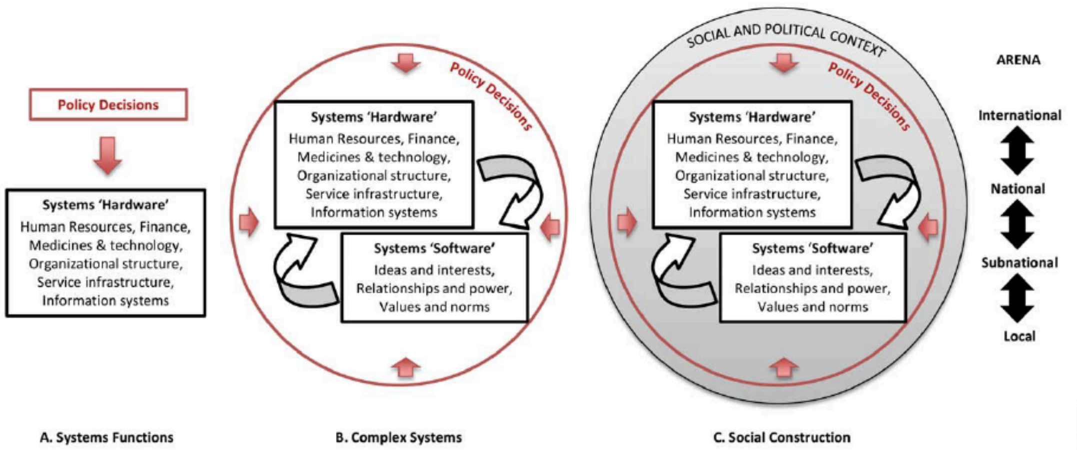
Frenk (2010), on the other hand, defines the health systems in terms of its stakeholders. He identifies patients, consumers, and tax-payers as key players in his health system framework and outlines how the health system operates through these stakeholders as the main sources of financing and co-producers of health. Fulop et al. (2001) and Van Damme et al. (2010) characterize the health system in terms of its level of operation. They identify three different levels - macro, meso, or micro. As per their framework, macro includes the national and international context which determines policy such as resource allocation and financing policy, meso involves the local health system and/or the organizational level which determines how policy gets implemented, and micro involves the people in the system (both patients and providers).

Another class of macro-systems frameworks are more analytical in their objective, with an additional focus on specifying relationships within the system. These frameworks tend to specify the form of the relationship between system elements either in a generalized way (through indicating which elements of the system are interlinked or not linked) or in a specific way by drawing on theories to define the nature of relationships. For example, Frenk (1994) identifies the state, health providers, and the population as key components within a health system, with a series of arrows showing how these actors are linked. He argues that the relationship between providers and the population does not occur in isolation but is rather shaped by the organizations in which they operate, the heterogeneous nature of the organizations and the population, and the state through setting policies of regulation and financing. While the author theorizes how and why these relationships exist, he does not employ specific theories to explain the nature of these relationships.

On the other hand, Gilson (2003) characterizes the health system through relationships between patients and providers, and defines the relationship through very specific 'trust' flows. She argues that the behaviours of health system providers and patients are directly influenced by trust between the patient and the provider, and trust between the health agent and the wider institution.

The work of both Frenk (1994) and Gilson (2003) points towards the need for considering the software (i.e. institutional environment, values, culture and norms) in addition to the hardware (i.e. population, providers, organizations) of a health system in order to understand the dynamics at play. In line with their arguments, Sheikh et al. (2011) propose a model which builds in such software into the existing 'building blocks' of health systems as proposed by the WHO (2007). They argue that in addition to the WHO building blocks, the values, norms, ideas, and power dynamics play a critical role in how relationships between different system elements are shaped.

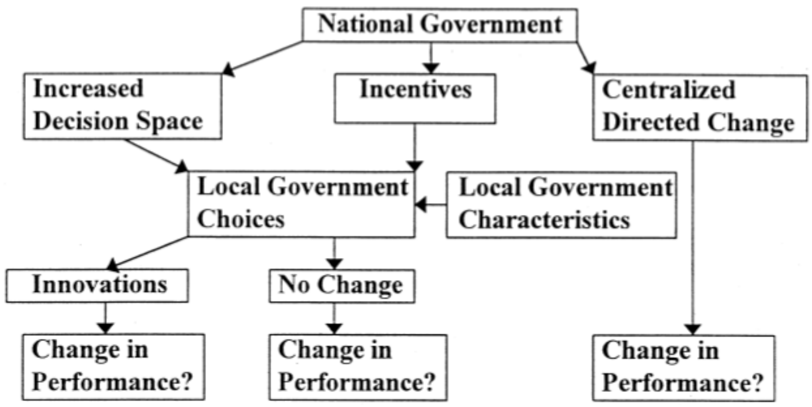
Figure 2: Health System Hardware and Software



Source: Sheikh et al (2011)

Often frameworks tend to focus on a sub-system to characterize relationships between different system components. Kutzin (2000), for example, develops a macro-systems approach focused on health financing. He outlines the various financing system functions in health financing including revenue collection, pooling of funds, purchasing of services, and provision of services to identify specific policy levers for the government to improve access to health finance. Similarly, Bossert (1998) develops a framework to study decentralization in health systems across countries. He uses the principle-agent model as his basis and extends the model through what he calls the 'decision space approach' to understand the degree of choice transferred from the center to local authorities and the impact of this choice on performance.

Figure 3: A Decision-Space Approach for Studying Decentralization

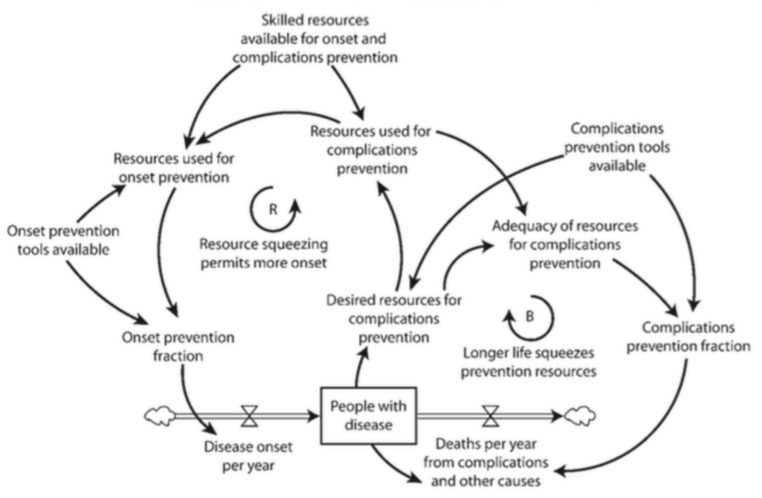


Source: Bossert (1998)

A final set of macro-system approaches use very specific numerical values to characterize relationships between system components, such as through the use of systems modelling, systems dynamics, and causal loops. While research in formal modeling of health systems continues to grow, the effectiveness of such models tends to be limited to their predictive ability instead of being able to study the actual impact of large-scale interventions (Hanson 2015).

For example, Homer and Hirsch (2006) develop a causal diagram of how chronic disease prevention works and then use systems dynamic methodology (grounded in concepts of accumulation and feedback loops) to develop a computer-based model to test alternate policy scenarios that may affect the chronic disease population (see Figure 4). Rwashana et al. (2009) use dynamic synthesis methodology (DSM) to model sub-systems within the immunization system (parental participation sub-system and healthcare sub-system). They use this model to explain uptake of immunization in Uganda. Another example is Batterham et al. (2002) where the authors use concept mapping to understand GP integration across primary and secondary health care systems in Australia. They develop a typology and model of GP integration using concept mapping in 11 groups of GPs, consumers, and other practitioners and then test it through confirmatory factor analysis. Bishai et al. (2014) explore how a hypothetical policy change of funding curative versus preventative services might lead to unintended consequences through complex relationships between stakeholders and financial resources. They identify several negative feedback loops that lead to stable model equilibria that were unexpected from the objectives of the original policies.

Figure 4: A simple model of chronic disease prevention



Note. The rectangle represents a stock of people; thick arrows with valves and cloud symbols represent flows of people; thinner arrows indicate causal influence; arrows with minus signs indicate inverse causal influence.

Source: Homer and Hirsch (2006)

2.3. Micro-systems Approaches

Micro-systems approaches in health aim to answer questions with respect to how a specific policy is designed, implemented, evaluated, and scaled-up (Mills, 2012; Gilson, 2012; Hanson 2015), with a specific focus on how policies interact with other system components. The research draws on a range of disciplines and methodologies to address these questions. In this sub-section, we illustrate these approaches.

Given health systems researchers often deal with complex policies and interventions that interact with various elements of the system and context in different ways, a key question for HSR is how to think about research and evaluation design, measurement of outcomes, and process evaluation whilst accounting for these interactions (Hawe, 2015). To address these set of concerns, the field has seen a rise of 'realist evaluations' as a way to evaluate complex interventions to tease out causal relationships. This technique recognizes that many different variables may be interwoven which interact in different ways with the fabric of society. Hence, the aim is to identify 'what works in which circumstances and for whom?', rather than merely 'does it work'? (Pawson and Tilley, 1997). More specifically, instead of looking at simple cause and effect relationships, realist research considers the interaction between context (the specific setting in which an intervention is rolled out), the mechanism (process of how an intervention works) and outcome (C-M-O). It develops 'middle range theories' through developing context-mechanism-outcome relationships which show how an intervention works (Greenhalgh et al., 2016).

While such evaluations have been rising in the field, their uptake has been slow due to lack of clarity around the methodology, lack of guidance on its use, and its time-consuming nature (Marchal et al., 2012). Despite these challenges, some researchers have been able to leverage the methodology effectively to tease out important insights. For example, Mac Kenzie et al. (2009) use realist evaluation techniques along with a clustered randomization trial to understand the impact of a nutritional intervention during a smoking cessation programme. The authors argue that using realist approaches helped them build a more refined understanding of how outcomes and processes were related. While Mac Kenzie et al. (2009) combine realist approaches with a rigorous experimental design, a lively debate continues on whether realist approaches can be used together with experimental and quasi-experimental approaches. While proponents argue that realist approaches can be integrated with RCTs by focusing on standardizing processes and functions of interventions (Hawe et al., 2004; Bonell et al., 2012), others argue that given RCTS are fundamentally based on a positivist paradigm, they would be unable to fully adapt to capture the complexity of interactions.

In addition to realist evaluations, evidence aggregation techniques such as meta-analysis and systematic reviews are also used in HSR. Leviton (2017), for example, argues that systematic reviews and meta-analyses can offer bodies of knowledge that support better understanding of external validity by identifying features of program theory that are consistent across contexts. To identify these systematically, she identifies several techniques to be used in combination with meta-analyses such as a more through description of interventions and their contexts, nuanced theories behind the interventions, and consultation with practitioners.

Health systems researchers are also beginning to rely on an evidence aggregation method called *realist synthesis*, which relies on the realist philosophy. The key idea is to aggregate evidence along the context, mechanism, outcomes outline (C-M-O) to identify not only the average treatment effect, but also how an intervention was intended to work (Wong et al., 2013). Abimbola et al. (2019), for example, carries out a realist synthesis of decentralization interventions to understand why, how, and in what context decentralization effects health system equity, efficiency and resilience. The author identifies three mechanisms which may mediate the effect of decentralization on health outcomes: 1) 'Voting with feet' which captures how decentralization affects patterns of inequities in a jurisdiction; 2) 'close to ground' which captures how local governance allows for local initiative, input, feedback; and 3) 'Watching the watchers' which captures the mutual accountability links between the citizens and the government. Greenhalgh et al. (2016) conduct a realist review to understand how community pharmacies support smoking cessation. Their review identifies five mechanisms that could support success or failure of pharmacy-led smoking cessation programmes - pharmacist identity, pharmacist capability, pharmacist motivation, clinician confidence, and public trust.

Understanding gaps in policy implementation is a key focus of health systems research (WHO 2002; De Savingy and Adam 2009). For example, Sheikh and Porter (2010) conduct a stakeholder analysis to identify key gaps in policy implementation. They use data from 46 in-depth interviews with various stakeholders across 5 states India to understand bottlenecks in HIV policy implementation (from 9 hospitals selected by principles of maximum variation). Using the "framework" approach for applied policy analysis, combining inductive and deductive approaches, they find that key gaps in policy implementation included conflicts between different actors' ideals of performance of core tasks and conformance with policy, and problems in communicating policy ideas across key actors involved in implementation.

Another method that health systems researchers rely on is ethnography and participant observations, especially when the question of inquiry involves complicated relationships between different system actors and elements. For example, accountability relationships between different health system actors are central to health service delivery, but capturing the complex social and political realities around such relationships requires techniques which allow deeper exploration. For example, George (2009) conducts an ethnographic analysis to understand how social dynamics may create individuals own meaning of accountability. He examines routine human resource management and accountability practices in Koppal state, India, showing how a complex web of social and political relations among different actors in primary health care influences local understandings and channels of accountability.

Systems research often has a specific focus on the implementation, uptake, and scale-up of policy (Hanson, 2015). The discipline of *implementation science* in the health sector is specifically targeted towards understanding such issues (Rubenstein and Pugh, 2006). The discipline has been defined as "the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services" (Eccles and Mittman, 2006). Greenhalgh et al. (2017), for example, combine qualitative interviews, ethnographic research, and systematic review to study the implementation of technological innovations in health. They develop the non-adoption, abandonment, scale-up, spread, and sustainability (NASSS) framework to both theorize and evaluate the implementation of health care technologies. Research in implementation science is at times less concerned with the question of what is effective (where there is strong prior evidence on an intervention's efficacy in ideal conditions) and is more concerned with how to implement effectively. For example, there is a comparatively large body of research on evidence-based treatments in mental health services, but their adoption and implementation in practice remains a challenge (Procter et al., 2009).

Discrete choice experiments are a methodology that have been adopted by systems researchers to understand questions such as patient preferences for different aspects of a treatments and health worker job preferences (Ryan, 2009). This has enabled researchers to develop clarity on these questions, within a single context as well across several contexts, in a cost-effective way. For example, Blaauw et al. (2010) use discrete choice experiments (DCEs) to evaluate the effectiveness of different policies in attracting nurses to rural areas in Kenya, South Africa and Thailand. They find that in Kenya and South Africa, better educational opportunities or rural allowances would be most effective in increasing the uptake of rural posts, while in Thailand better health insurance coverage would have the greatest impact. Such approaches are also helpful in developing a system-wide understanding of central questions such as job preferences which is essential for understanding how to allocate limited resources to achieve health gains.

In response to the complexity in systems and interventions, some systems researchers rely on methods that allow for more iterative experimentation and learning. For example, Tsofa et al. (2017) use a 'learning sites' approach in which a geographical space is specifically created where researchers and health system practitioners work together over long periods of time to uncover and address thorny governance challenges. As part of the learning site activities, formal reflective sessions are regularly held among researchers, between researchers and practitioners, and across learning sites to develop an in-depth contextual grounding to study complicated pathways to change. Using this approach, the authors study the impact of a new decentralization reform in Kenya on health resource allocation and budgeting. They conclude that the decision space, organizational capacity, and accountability structures are critical to achieving decentralization success.

3. Education Sector

3.1 Motivation, Definition, and Scope

Education systems research emerged with the growing recognition that significant investments in various inputs such as textbooks, hiring of new teachers, and increased salaries of teachers have not had the effects that governments and researchers hoped for in terms of improvements in learning outcomes (Banathy, 1991; Betts, 1992; Glewwe and Muralidharan, 2015; World Development Report, 2018).[3] In some of the early works on education systems research, Banathy (1991) argued that a new systems framework for creating educational change was needed in light of the changing demands of societies and the needs of various stakeholders within a system. More recently researchers and practitioners are increasingly recognizing that the current global 'learning crisis' requires addressing system weaknesses and making the whole education system coherent with learning (Pritchett 2015; World Development Report, 2018). This has brought a greater emphasis on understanding the interdependencies between various features of an education system such as institutions for governance, accountability, information, financing rules, and school management (World Bank, 2014; World Development Report, 2018).

Moore (2015, p. 1) defines education systems as "institutions, actions and processes that affect the 'educational status' of citizens in the short and long run." In line with some of the early calls to adopt a systems lens for understanding education reform (Banathy 1991; Betts, 1992), several global institutions are making strides towards undertaking education systems research. For example, a World Bank initiative called Systems Approach for Better Education Results (SABER) was launched in 2011 with the goal to collect data on system capacities and gaps

[3] For example, Indonesia doubled teacher wages incurring an expenditure of nearly 4.5 billion USD which produced near-zero impact (De Ree *et al.,* 2015). Similarly, research from India and Africa shows that reduction in class-sizes does not always produce the desired impact if other systemic features such as teachers, incentive structures, and curriculum do not change (Pritchett, 2015).

through a range of survey tools designed for each education sub-system. The Research on Improving Systems of Education (RISE) programme is another example of a research program focused on education systems research. RISE is a multi-country programme that was initiated in 2015 with research in Pakistan, Ethiopia, India, Tanzania, Indonesia, and Vietnam. The programme aims to conduct empirically and theoretically well-founded interdisciplinary research to understand how education systems function. Another example of a research programme that takes a systems approach is the Raising Learning Outcomes in Education Systems (RLO) research programme funded by the Foreign, Commonwealth & Development Office (FCDO) and the Economic and Social Research Council (ESRC). The programme focuses on interactions between different education system components and the various contextual features that affect educational reform.

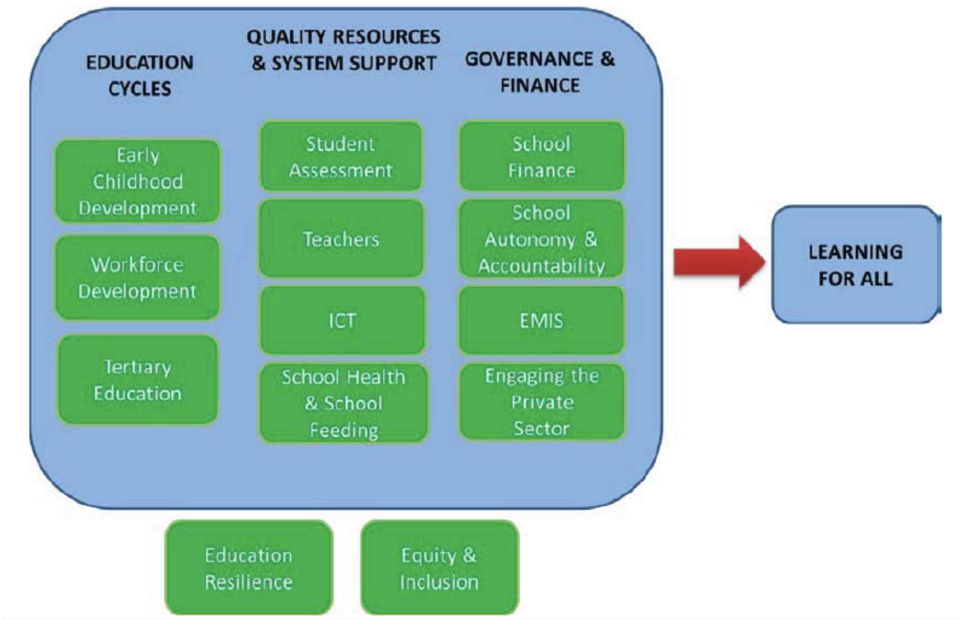
The scope of education systems research has close parallels with health. Pritchett (2015b) and Hanson (2015) outline the following key features of education systems research. First, it aims to be 'question-driven' which employs various disciplines as per the need of the question. For example, systems research in education could explore questions ranging from the impact of a national teacher training intervention through mixed methods to questions about how power and accountability structures in the education system function through ethnographic accounts drawing on various disciplines such as economics, political science, sociology, and anthropology. Second, the questions focus on real-world issues which either pertain to the system as a whole or a specific sub-component. For example, while questions about national teacher recruitment or training policy would be relevant for the teaching system sub-component, questions about a specific teacher training programme in 10 selected villages by a specific NGO would not be relevant as they would not have any implications for the teaching system or the education system as a whole. Third, systems research in education aims to explore questions which relate to learning gains. Fourth, it studies reforms which have the potential for scale and fit the context of the specific region.

3.2 Macro-systems Approaches

Similar to health systems research, macro-systems approaches in education also vary in the specificity with which they define relationships between system components. In this sub-section, we outline a range of macro-systems approaches in education systems research.

Some macro-system approaches define different components within an education system, without specifying relationships between different sub-components. Such approaches in education systems research have often been used as a foundation for designing survey tools for system diagnostics. For example, Systems Approaches for Better Education Results (SABER) at the World Bank describes the education system in terms of 13 different functions (Halsey and Demas, 2013).

Figure 5: Domains of Education System (SABER)



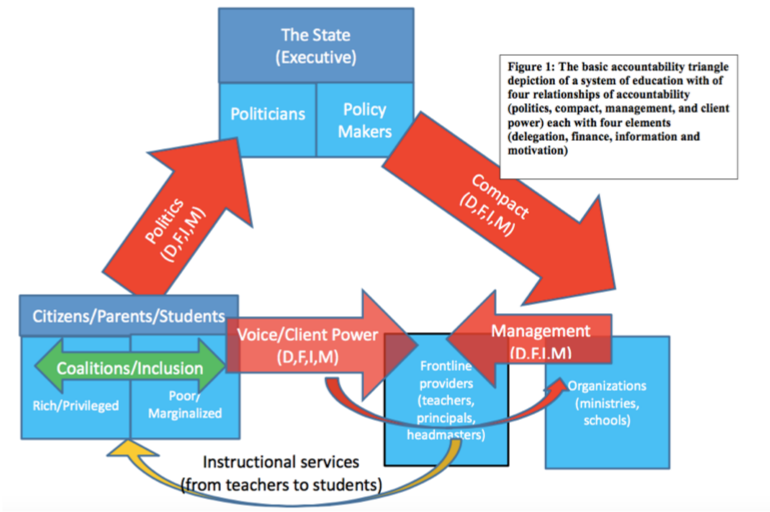
Source: Halsey and Demas (2013)

SABER has implemented its system diagnostic tools in more than 100 countries to identify key constraints to system effectiveness (World Bank, 2014). In Jamaica the government's Early Childhood Commission employed the SABER dataset to draft its new National Strategic Plan as well as a national multi-sector early childhood development policy. Similarly, in Tanzania, information from the diagnostic supported the Government to plan its education reforms (World Bank, 2014). Country teams in RISE have also adapted these tools locally to develop clarity on how various sub-components in the education system contribute to (or hamper) system effectiveness. Pritchett (2018) highlights key insights and challenges of using such system diagnostics. He argues that input indicators and de jure (formal) policies which these tools aim to capture do not always explain learning - for example, Vietnam shows high learning scores in PISA assessments but indicators in the system diagnostic tools are unable to explain this success. He argues that to understand the drivers the system effectiveness, it is essential to develop and implement tools which in fact aim to capture *de facto* (i.e. actual) policies. At the same time, he acknowledges that developing and implementing tools that capture *de facto* policies can be challenging.

Another macro-systems framework which is descriptive in nature is the General Education Quality Analysis Framework (GEQAF). This framework defines 5 components of the education system, with each component catering to a specific impediment to learning: 1) supporting mechanisms (which includes governance, financing and system efficiency); 2) core resources (which includes curricula, learners, teachers and the learning environment); 3) core processes (which includes learning, teaching and assessment); 4) desired outcomes (which includes competencies and life-long learning; and 5) development goals (which includes relevance and equity) (UNESCO, 2012). While these frameworks go in a fair degree of descriptive detail, they do not specify how different components of the education system are related to one another.

Some macro-systems approaches in education systems research specify relationships between system components. For example, Pritchett (2015) describes the education system components and the relationships between them through very specific accountability relationships. He describes the education systems as a composition of the following actors - the executive apparatus of the state which makes key decisions (laws, regulations, policies and the allocation of budgets); organizational providers of schooling such as schools and organizations that control and manage the schools; teachers who are the "front-line service providers"; and citizens/parents/students who are the intended beneficiaries of schooling. He defines the relationships between these different actors as 'accountability' links which act through four design elements– delegation, financing, information, and motivation. He argues that the system of education works when there is an adequate flow of accountability across the key actors in the system across these four design elements (see Figure 6).

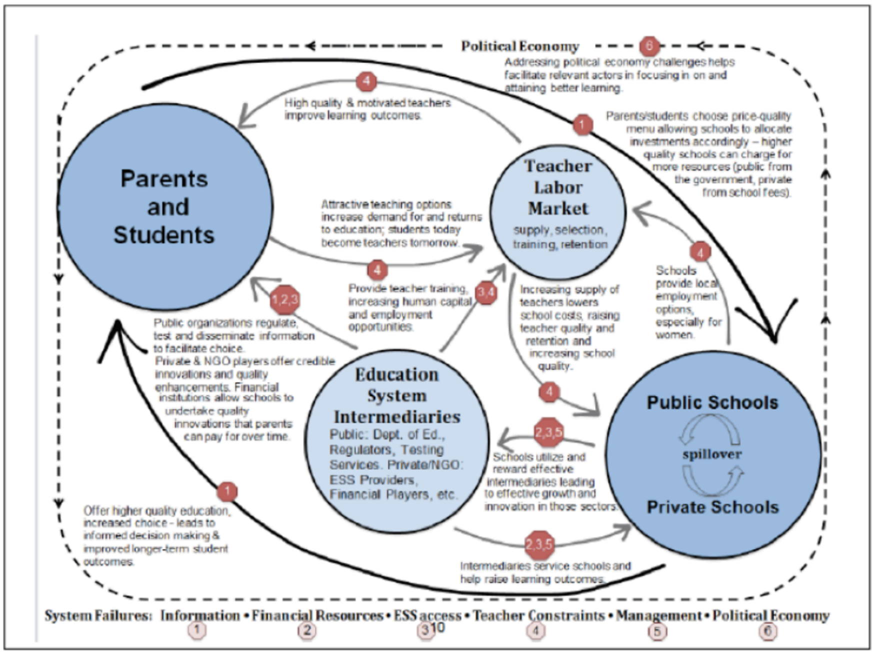
Figure 6: Accountability Triangle in the Education System



Source: Pritchett (2015)

Similar to Pritchett (2015), Andrabi et al. (N.D.) also describe the relationships within the education system through specific relationships. They describe the Pakistan education system as an economic market with key constraints and frictions along five dimensions: 1) access to information; 2) access to resources and financing; 3) knowledge and innovation markets; 4) labor market incentives; and 5) regulatory and governance structure. They argue that the functioning of the education system hinges on being able to address key frictions in the market along the above-mentioned dimensions (see Figure 7 below).

Figure 7: Market Frictions in the Education System



Source: RISE Pakistan Technical Proposal

Similar to the health sector, education systems researchers are also starting to model education systems where relationships between system components are defined through very specific numerical relationships. For example, Kaffenberger and Pritchett (2021) develop a structural model to capture the dynamics of learning. Using existing empirical literature to assign numerical values to the parameters in their model, they predict how learning outcomes would be affected under different policy scenarios such as expanding schooling to universal basic education, slowing the pace of curriculum, and increasing instructional quality.

3.3 Micro-systems Approaches

Micro-systems approaches in education systems research are characterized by a focus on not only what works, but also how and why (Magrath et al., 2019). Given this focus, they tend to rely on a combination of quantitative and qualitative methods to understand whether, how, and why specific policies work, often comparing their impacts across contexts.

Magrath et al. (2019) highlights several projects under the Raising Learning Outcomes in Education Systems (RLO) research programme funded by FCDO and the ESRC where researchers are using mixed-method approaches to diagnose questions such as how pedagogical reform takes place or how accountability relationships function. For example, Lynch et al. (2018) use a mixed-methods study following the guidelines from the the Medical Research Council (MRC) Framework for Developing and Evaluating Complex Interventions (Craig et al., 2008) to design and test the feasibility of a training programme for developmental stimulation of children with visual impairment in Malawi. Using qualitative interviews to guide the initial training design, the researchers used a combination of quantitative data from logbooks along with in-depth interviews to assess the fidelity of implementation, as well as insight into outcomes that needed to be measured to understand the impact of such training programmes. In another example, Aiyar et al. (2015) combine qualitative interviews with quantitative time-use data of the public education officials in Bihar to understand how a new pedagogical reform works. Their study highlights that organizational culture plays in instrumental role in how reforms are perceived and implemented by frontline workers

Over the last two decades, the education literature has also explored a series of system-level questions through experimental and quasi-experimental techniques. Although surveying the full range of these studies is outside the scope of this review, they include understanding the impact of large spending by governments (in the form of textbooks, cash transfers), governance reforms such as teacher incentives or community monitoring programmes, new pedagogical approaches in government schools (such as contract teachers or literacy and numeracy skills lessons), and the impact of school-based management reforms (see Glewwe and Muralidharan, 2015 for a review of studies). Many of these experimental studies design multi-arm-controlled trials to look for interactions between arms of a policy and other features of policy context in an effort to understand theoretical mechanisms. For example, Andrabi et al. (2018) design a multi-arm intervention to test the impact of alleviating financial constraints for private schools. The variations in treatment arms, by providing cash transfers to either one private school in the village or all private schools in the village, allow the researchers to understand how financial constraints interact with the overall market structure.

With the surge of experimental and quasi-experimental studies in the education sector, there has also been a focus on methods of evidence aggregation such as meta-analysis and systematic reviews (Conn, 2014; McEwan, 2015; Evans and Popova, 2016 to name a few). These reviews, which have largely focused on identifying average treatment effects of interventions across contexts, have also at times pointed to concerns of external validity and how similar interventions can have very different effects across contexts or when scaled up (Pritchett and Sandefur 2015; Bold et al., 2016; Masset, 2019). For example, Masset (2019) calculates prediction intervals for various meta-analyses of education interventions and finds that interventions' effectiveness is highly heterogeneous and unpredictable across contexts, even for simple interventions like merit-based scholarships.

Similar to the health sector, education systems researchers also tend to rely on the methodology of realist synthesis to understand how interventions work. For example, Eddy-Spicer et al.

(2016) conduct a realist synthesis to understand how school accountability policies (such as assessments, monitoring, and inspections) operate locally in schools in low- and middle-income countries to improve student learning outcomes. The findings highlight that improved student learning outcomes tend to be associated with stronger support structures for school leaders and staff in how accountability policies are implemented.

One type of qualitative methodology used by education systems researcher is ethnography and participant observation. For example, Bano and Oberoi (2020) use ethnographic methods to understand how innovations are adopted in the context of an Indian NGO that introduced a Teaching at the Right Level (TaRL) intervention, and tease out lessons for how innovations can be scaled and adopted in state systems. Watkins and Ashforth (2019) aim to understand norms and practices around schooling by observing interactions between parents, teachers, and administrators in rural Malawi at the grassroot level. Using narratives from the Malawi Journals Projects, interviews in the study, and participant observations, the researchers highlight how issues of accountability at the school-level are resolved on a daily basis.[4]

Education systems researchers often rely on large longitudinal quantitative datasets to answer questions about specific policies. For example, Young Lives is a longitudinal study of 12,000 students across the countries of Ethiopia, India, Peru and Vietnam covering a life span of 15 years. This project has created a rich longitudinal household and student learning dataset overtime which has allowed the team to explore important policy questions with a systems lens. For example, in India the household surveys and learning data have together shed light on the role played by low-cost private schools within the education system.[5]

---

[4] The Malawi Journals project is an account of narratives written by individuals from rural Malawi from 1999-2015. See details here: https://deepblue.lib.umich.edu/handle/2027.42/113269

[5] https://www.younglives.org.uk/content/education

Practitioners and researchers in education are also increasingly focusing on complexity in education systems, and the set of strategies that may be required to account for this complexity when designing educational reform (Snyder, 2013; Crouch and Destafano, 2017). Snyder (2013) draws on complexity theory and its applications to health and ecology, and argues how principles of complexity theory can be applied to educational reform. Crouch and DeStefano (2017) propose a strategy for intervention design and evaluation that relies on the strategies of 'Doing Development Differently (DDD)' which rely on local-level problem identification and problem-solving, involving a process of iteration and adaptation.[6]

Within the education sector, several research efforts try to understand the management, governance, and performance of education systems at a more macro-level through the use of micro data. These research efforts are hard to classify into our macro and micro systems research classification, but serve as unique examples of research that tries to bridge these two types of research. For example, education system researchers often rely on large longitudinal quantitative datasets that rely on micro data to answer questions about how education systems function. For example, Young Lives is a longitudinal study of 12,000 students across the countries of Ethiopia, India, Peru and Vietnam covering a life span of 15 years. This project has created a rich longitudinal household and student learning dataset overtime which has allowed the team to explore important policy questions with a systems lens. For example, in India the household surveys and learning data have together shed light on the role played by low-cost private schools within the education system (Rossiter et al., 2018).

[6] The DDD strategies relate closely the 'Problem Driven Iterative Adaptation' approach proposed by Andrews et al (2013)

On the other hand, Adelman et al. (2021) have developed a new instrument called the Education System Coherence Survey (ESCS) that aims to understand coherence in the understanding of task allocation of bureaucrats across the education delivery chain. One of the measures that can be constructed from the instrument is an incoherence index that captures the gap between *de jure* task allocation and bureaucrats' *de facto* understanding of task allocation, allowing researchers to explore how such incoherence at different levels of the education system may be related to student learning outcomes. This survey has been implemented across four LAC countries—Brazil, the Dominican Republic, Guatemala, and Peru. Another example is the work of Levy et al. (2018) that explores the performance of education systems as a whole, but with a focus on politics and institutions. The authors present a multi-level framework where incentives and constraints at the national level shape incentives and constraints at various sub-national levels of the education bureaucracy, ultimately cascading down to schools. Applying the framework to the context of two South African provinces, where there is significant delegation of education service delivery to provinces and schools but variation in provincial-level political dynamics, they explore conditions under which horizontal and/or hierarchical models of governance work.

4. Infrastructure Sector

4.1 Motivation, Definition, and Scope

Systems research in infrastructure is conducted with the primary aim to understand and manage complex interactions within and between various infrastructure sectors. Research in this sector can be clearly demarcated into two categories where each has its own motivation and objective: 1) sector-specific system analyses which allows taking a systems approach within a specific type of infrastructure sector (e.g. water, electricity, or gas); and 2) systems-of-systems analyses where research is conducted across various infrastructure sectors to explore relationships between infrastructures sectors, infrastructure risk, and long-term systems-of-systems analyses. While the former is considerably well-established, the field of system-of-systems analyses is relatively new (Hall et al. 2016).

Sector-specific system analyses are motivated by the idea that specific infrastructure sectors can be made more efficient by understanding feedbacks between various system components, and managing their demand accordingly. A systems-of-systems approach takes account of the cross-sectorial interdependencies between different infrastructure sectors and is motivated by two key needs that infrastructure systems face today. First, the need for adequate planning for future operation, capacity, and environmental performance of infrastructure systems in light of future socio-economic changes such as population changes, per-capita infrastructure demands, and economic growth. Second, the need to ensure resilient operation of infrastructure services in the face of increasing climate and socio-economic risks. These challenges are exacerbated by the fact that infrastructure networks have become increasingly interdependent, providing potentials for knock-on effects causing major economic and societal disruptions. For example, a power failure in a major electricity exchange can result in the temporary loss of broadband service for hundred thousand of households and businesses (BBC 2011). Hence, systems-of-systems approaches for short-term risk analysis aim to reduce the risk of cascading infrastructure failures, allow for more effective responses, and improved coordination (Dudenhoeffer et al., 2006).

A specific infrastructure system can be defined in several ways depending on the type of infrastructure or the scope of research analysis. While generally infrastructure systems are understood as various interdependent physical and socio-economic systems to distribute essential services, (Bissell, 2010), another approach to define infrastructure systems is through the types of assets within the system which can include energy, transport, water, waste, information and communications technology (ICT), social infrastructures (hospitals, schools, etc.), financial services, and the built environment (Cabinet Office, 2010). Analysis of infrastructure at a system level requires integration of various components – such as across different scales (e.g. urban, rural, or regional), across eco-systems (e.g. social, urban, land, water and climate), and between different structures or sectors (e.g. social, physical, health, economic and political). Following this, Hall et al. (2016, p.6) develop a definition of infrastructure systems as 'the collection and interconnection of all physical facilities and human systems that are operated in a coordinated way to provide a particular infrastructure service."

Infrastructure systems research is used to understand current infrastructure performance (for example, whether different infrastructure sectors currently meet demand, environmental standards, resilience criteria), predict future infrastructure needs, and to understand the impact of newly built infrastructure assets on the entire system. The scope of systems research in infrastructure has several commonalities with the health and education sectors. First, it has a real-world focus, where approaches and methodologies for system assessments of infrastructure are direct real-world problems of planning, designing, and operating infrastructure. Second, it tends to be multi-disciplinary. While the bulk of analysis in the field includes quantitative modelling, it often combines qualitative approaches such as simulation modelling with decision science, policy and governance research, and adaptive pathways. Third, it focuses on directly impacting policy. For example, a number of infrastructure assessment methodologies inherently include adaptive pathways and policy recommendations.

4.2 Macro-systems Approaches

Macro-systems approaches in infrastructure systems research help policymakers and researchers understand how the entire system functions. This helps answer questions such as whether governments should make a large investment in an infrastructure asset or how to manage risks of infrastructure failures. Given these are high stake concerns for governments, macro-systems approaches in the infrastructure sector mostly include models with tightly specified numerical relationships that can make accurate predictions.

These tightly specified models tend to characterize relationships between different system components through what the infrastructure sector calls *'interdependencies'*. Researchers have adopted descriptive approaches to identify a range of such interdependencies. For example, Rinaldi et al. (2001) outline that interdependencies depend on the scope of the framework and can be classified as a) physical (material or physical flow from one entity to another); b) cyber (information transfer); c) geographical/spatial (physical proximity affecting components across multiple infrastructure systems); or d) logical (dependencies other than the above three categories). Dudenhoeffer et al. (2006) further expand these classes to include two additional categories: a) policy/procedural which includes the effect of a policy or a procedure of one infrastructure on all other social and economic sectors; and b) societal which captures the effect of all influencing factors such as public opinion, confidence, fear, or cultural issues from one system component to another. These different types of interdependencies tend to form the basis of how relationships between different system components within an infrastructure system are characterized.

To developed models of infrastructure systems (that rely on these interdependencies), the literature proposes several different approaches (Ouyang, 2014; Saidi et al., 2018; Dudenhoeffer et al., 2006; Xiao et al., 2008), that can be classified into five broad categories - system dynamic-based approaches, agent-based simulation and modelling, input-output models, network-based approaches, and empirical approaches. A growing number of studies suggest that the current infrastructure system is most suitably modelled using a network-based approach of nodes and edges, which capture essential interdependencies and indicate the flow of directionality across infrastructure assets (Lewis, 2006). While no network modelling approach can answer all the questions (Brown et al., 2004; Eusgeld and Kroger, 2008), models which incorporate systems theory and develop networks which adapt to their environment are considered to be the state-of-the-art (Eusgeld and Kroger., 2008; Xiao et al., 2008; Ouyang, 2014; Bevir, 2007).

For example, Dudenhoeffer et al. (2006) use a conventional graph theory concept to define an infrastructure system as a collection of nodes, links, and edges which represent the dynamic and complex nature of the system. The dynamic aspect of the system is demonstrated by the fact that the network can grow overtime (through increase in the number of nodes); it can evolve (through changing links between the nodes), or entail complexity (through non-linear effects of nodes on one another which also change the state of the nodes). Saidi et al. (2018) develop a similar multi-layered framework for the civil infrastructure system (see Figure 8) which shows different types of interdependencies between various physical infrastructure sectors, and the broader social, economic, and political environments. The framework also clearly identifies the type of relationship as physical, geographical, logical, or cyber. Such multi-layered networks offer a type of 'systems-of-systems' framework which model a range of interdependencies across different infrastructure sectors.

Figure 8: A systems-of-systems view with different dependency types

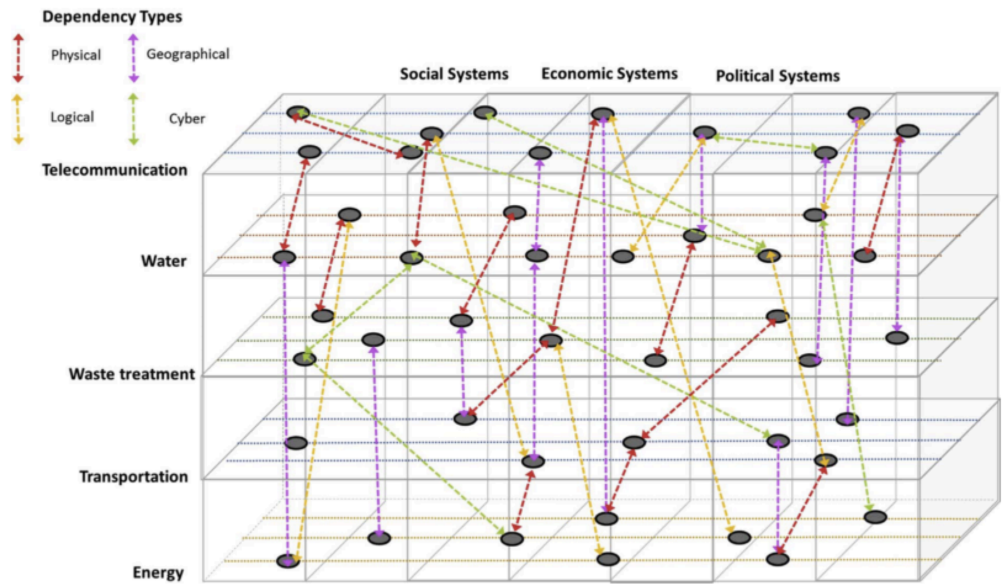
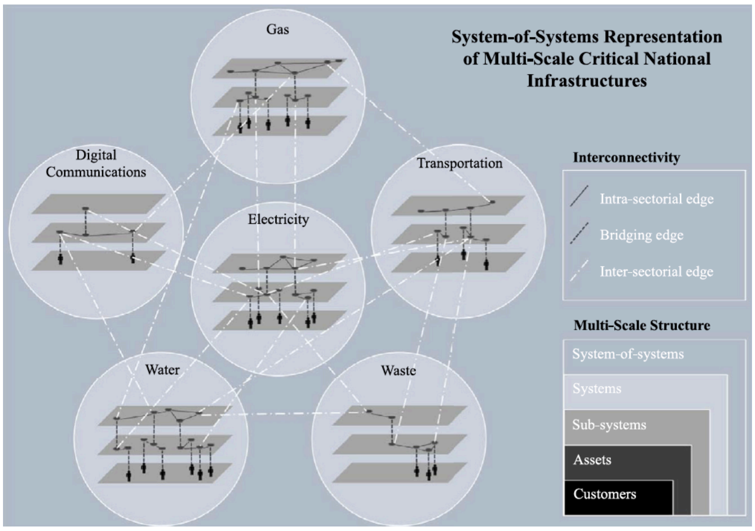


Fig. 1. Multilayer network of an integrated civil infrastructure system with different dependency type.

Source: Saidi et al. (2018)

The level of detail with which such relationships/interdependencies are specified varies, and primarily depends on the goal of the framework. For example, disruption analysis inherently involves detailed specification of interdependencies at the outset of the analysis whereas predicting long-term performance of infrastructure systems may not involve the same level of detail. Thacker et al. (2017), is an example of the former. The authors characterize critical national infrastructures as a system-of-system framework to perform a multi-scale disruption analysis. Their framework requires a detailed specification of the physical and geographic network interdependencies between sectors. The authors model each type of infrastructure such as water or electricity as a sub-system comprising of a group of nodes and edges with their specific flows (see Figure 9). They use this model to perform a multi-scale disruption analysis and draw predictions on how failures in any individual sub-systems can potentially lead to large disruptions.

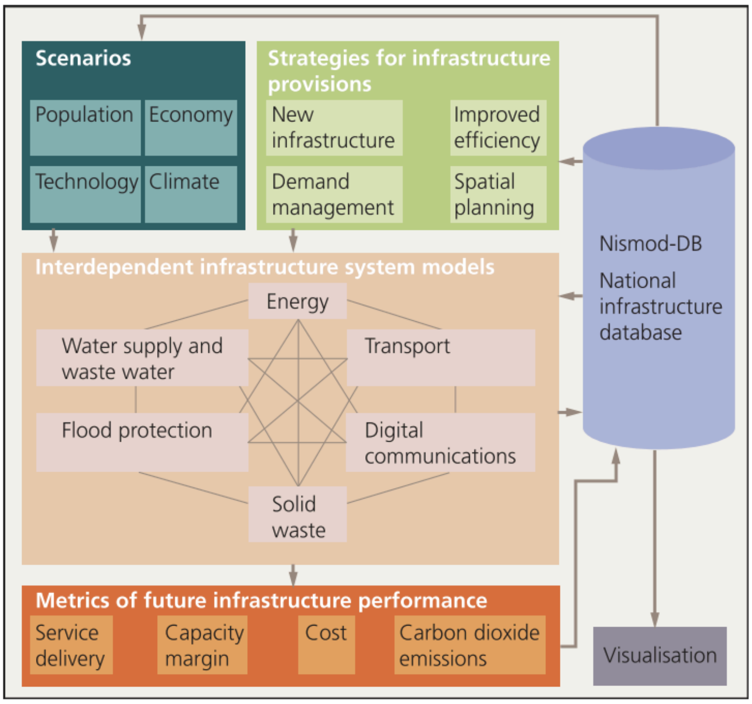
Figure 9: System-of-systems representation of six critical national infrastructures



Source: Thacker et al (2017)

On the other hand, Hall et al. (2017) develop a national infrastructure assessment framework with the aim to assist decision-makers in analyzing the long-term performance of interdependent infrastructure systems. In contract to Thacker et al. (2017), this requires less detailed infrastructure interdependency modelling and a greater focus on understanding the common underlying drivers for infrastructure demands across sectors. This framework deals with each infrastructure sector – energy, transport, digital communications, water supply, waste water, flood protection, and solid waste – in a consistent model and assesses exogenous socio-economic drivers which may impact on all sectors (e.g. population growth, the rise of integrated ICT systems changes the demand patterns for classical infrastructures). It focuses on explicitly specifying how one sector may place demands on other sectors, or how a sector-specific capacity installation (a waste to energy plant) may add capacity in another sector (to electricity production). The focus of this framework on the national, long-term, and capacity/demand perspective leads to a choice of a comparatively descriptive system, because a very detailed representation of the interdependencies would be overconfident, over-complex and consequently, unhelpful (Otto et al., 2016).

Figure 10: Modelling Future Performance



Source: Hall et al (2017)

Some macro-systems frameworks do not focus on specific physical interdependencies but rather focus on explaining how a specific infrastructure project is influenced by its users, external stakeholders, asset managers, operators, and political decision-makers. Such frameworks, which include humans, often draw on qualitative disciplines to explain interdependencies. For example, Masood et al. (2016) develop a conceptual framework with the aim of future-proofing (i.e. anticipating future changes and needs to prepare appropriately to minimise ecological impact) infrastructure with two dimensions: infrastructure resilience (resilience to unexpected events) and change management capability (capability to adapt to changing needs). Ottens et al. (2006) propose a high-level framework to characterize how technical elements in an infrastructure system may interact with human actors and social institutions to determine system performance.

The focus and scope of macro-systems approaches used in infrastructure systems also depends on how system boundaries are drawn. For example, many authors view infrastructure services along with their management as part of the infrastructure system. The waste sector is one such example which includes both physical as well as management elements such as manufacturing, transportation, urban growth, development, land use, and public health considerations. This highlights the complexity between the physical components of the system and its social and environmental spheres (Seadon, 2010).

4.3 Micro-systems Approaches

Compared to health and education, assessing the impact of an infrastructure investment or a specific policy after its implementation is less common and hence the literature on the subject is less widespread. This can be attributed to the nature of infrastructure - its long-lifetime and costly resources warrant investment into detailed modelling to simulate how different infrastructure investments will perform in the future exante, with lesser focus on estimating the impact of the investment once it has been made.

One exemption to this is the development literature, where the effectiveness of an infrastructure intervention is often dependent on the local population using it. In such cases, impact of infrastructure is defined as how the infrastructure construction, rehabilitation or maintenance has affected people's lives (Hansen et al., 2011). The focus of development agencies on results and value for money has led to an increase in impact evaluations to demonstrate the effectiveness of infrastructure development programmes (Hansen et al., 2011). A range of quantitative methods are employed such as experimental methods (where random assignment is possible), quasi-experimental methods (in large-n cases), computational general equilibrium models (in small-n cases), and cross-country regressions. There has also been a recent surge in evaluating infrastructure investments for environmental outcomes, for example carbon emissions. Law et al. (2017) for example use energy analysis, an environmental accounting system, to evaluate the direct and indirect energy inputs into these infrastructures to give an indication of sustainability outcomes. Such infrastructure evaluations are valuable to decision makers and urban planners who aim to improve standard design and implementation practices for infrastructure projects.

Similar to health and education, infrastructure systems research can also rely on evidence aggregation methods such as meta-analysis to identify the impact of specific types of green infrastructures. For example, Filazzola et al. (2019) conduct a meta -analysis to study whether green infrastructure is beneficial for biodiversity as compared to conventional infrastructure.

Another area where micro-systems approaches are used in infrastructure systems research is in designing infrastructures. These approaches tend to be grounded in decision-analysis methodologies, which at times also draw on qualitative techniques. For example, scenario modelling and robust decision-making methods use multiple views of the future to identify conditions under which a decision would fail to meet its objectives (Lempert et al. 2006; Lempert et al. 2013). Similarly, hybrid methodologies tend to integrate stakeholder input into how infrastructure systems are designed. We give details on formal scenario planning, robust decision-making, and hybrid methodologies below.

Formal scenario planning embraces the concept of multiple future views (Bradfield et al., 2005). Scenarios are often presented as narratives of descriptions of possible paths into the future and can be differentiated in to three classes. These include probable scenarios (what will happen); possible scenarios (what could happen), and preferred scenarios (what should happen). Such scenarios are typically produced in group exercises where three to four such possible paths are generated (Wilkinson and Eidinow, 2008). These are intended as a set to stimulate group thinking and help decision-makers evaluate those strategies that perform well across multiple futures (Lempert et al., 2009). While it can be difficult to capture a wide range of potential futures in a limited set of scenarios, scenario analysis is the least complex of these techniques and has been widely employed for policy review and in infrastructure assessments.

Robust decision-making is applied using computer simulation models to test strategies against a range of potential futures. This involves considering hundreds to millions of scenarios – enough that one matches the actual future (Lempert, 2003). Such an exploration of the future aids policy-makers in determining those strategies in which performance is relatively insensitive, in other words 'robust', to key uncertainties. For example, Kalra et al. (2015) defined a robust portfolio of water reservoirs in order to implement Lima's long-term water resource plan. Such an approach can also help to define pathways that allow for flexibility and adjustment of the strategy once new information becomes available and future developments become more predictable.

Hybrid methodologies integrate stakeholders throughout the decision-making process for infrastructure development, prior or post modelling. Prior to the modelling, stakeholders may be engaged in defining which infrastructure interventions to model, or which criteria for performance modelling to choose (e.g. determining those infrastructure investments with least cost, least environmental impact, etc.). Such stakeholder methodologies typically make use of a number of methods, including Delphi or participatory backcasting. Delphi methods seek agreement on future infrastructure trends from a wide range of experts (Gordon, 1964). Such experts respond to a list of questions, review each other's answers, and revise their views accordingly in an iterative fashion. Stakeholders may further be integrated to define which infrastructure assets to model (e.g. building a new power plant, small solar parks, etc.) through participatory backcasting in which a single normative vision of the future is developed and different pathways are developed to reach that vision (Touminen et al., 2014). Tuominen et al (2014) propose a new strategy for backcasting studies called *pluralistic backcasting*, in which multiple visions of the future are developed through a participatory and interdisciplinary process that engages key stakeholders and users. Following this, policy packages that can potentially become pathways to these alternate visions are collaboratively developed with stakeholders. Post modelling, stakeholders can be integrated to encourage open discussion of trade-offs between different criteria, focusing on strategic, agreed-upon objectives rather than each stakeholder's personal cost and benefits.

References

Abimbola, Seye, Baatiema, Leonard, & Bigdeli, Maryam. (2019). The impacts of decentralization on health system equity, efficiency and resilience: A realist synthesis of the evidence. *Health Policy and Planning, 34*(8), 605-617.

Adelman, Melissa, and Renata Lemos. (2020). Managing for Learning: Measuring and Strengthening Education Management in Latin America and the Caribbean. Washington, DC: World Bank. doi:10.1596/978-1-4648-1463-1.

Andrabi. Tahir, Das, Jishnu, & Khwaja, Asim I. RISE in Pakistan: Technical research overview. https://riseprogramme.org/sites/default/files/publications/Pakistan%20CRT%20Technical%20Overview.pdf

Andrabi. Tahir, Das, Jishnu, Khwaja, Asim I, Ozyurt, Selcuk, & Singh, Niharika. (2020). Upping the Ante: The Equilibrium Effects of Unconditional Grants to Private Schools. *The American Economic Review, 110*(10), 3315.

Andreano R. 2000. *An informal assessment of the international health policy programme*. Geneva, Alliance for Health Policy and Systems Research. Working Paper.

Ang, Yuen Yuen. 2017. *How China Escaped the Poverty Trap*. Ithaca: Cornell University Press.

Banathy, B. H. (1991). *Systems Design of Education: A Journey to Create the Future*. Educational Technology.

Bandiera, Oriana, et al. 2016. *State Effectiveness, Growth and Development*. IGC, www.theigc.org/wp-content/uploads/2014/09/IGCEvidencePaperState.pdf.

Batterham, R, Southern, D, Appleby, N, Elsworth, G, Fabris, S, Dunt, D, & Young, D. (2002). Construction of a GP integration model. *Social Science & Medicine (1982), 54*(8), 1225-1241.

BBC. 2011. *BT suffers huge broadband failure*. http://www.bbc.co.uk/news/technology-15154020

Bett, Frank. (1992). How systems thinking applies to education. (Improving School Quality). *Educational Leadership, 50*(3), 38.

Bevir, M. (2007). *Encyclopedia of governance*. Thousand Oaks: Sage Publications, Inc.

Bissell, J. J. (2010). *Resilience of UK infrastructure*. *POSTNOTE*. Retrieved from http://www.parliament.uk/business/publications/research/briefing-papers/POST-PN-362

Blaauw, D, Erasmus, E, Pagaiya, N, Tangcharoensathein, V, Mullei, K, Mudhune, S, . . . Lagarde, M. (2010). Policy interventions that attract nurses to rural areas: A multicountry discrete choice experiment. *Bulletin of the World Health Organization, 88*(5), 350-356.

Blas, Erik, DrPH, Gilson, Lucy, Prof, Kelly, Michael P, Prof, Labonté, Ronald, Prof, Lapitan, Jostacio, PhD, Muntaner, Carles, Prof, . . . Vaghri, Ziba, PhD. (2008). Addressing social determinants of health inequities: What can the state and civil society do? *The Lancet (British Edition), 372*(9650), 1684-1689.

Bold, Tessa, Kimenyi, Mwangi, Mwabu, Germano, Ng'ang'a, Alice, & Sandefur, Justin. (2018). Experimental evidence on scaling up education reforms in Kenya. *Journal of Public Economics, 168*, 1-20.

Bonell, Chris, Fletcher, Adam, Morton, Matthew, Lorenc, Theo, & Moore, Laurence. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science & Medicine (1982), 75*(12), 2299-2306.

Bossert, Thomas. (1998). Analyzing the decentralization of health systems in developing countries: Decision space, innovation and performance. *Social Science & Medicine (1982), 47*(10), 1513-1527

Bradfield, Ron, Wright, George, Burt, George, Cairns, George, & Van Der Heijden, Kees. (2005). The origins and evolution of scenario techniques in long range business planning. *Futures: The Journal of Policy, Planning and Futures Studies, 37*(8), 795-812.

Brown, Theresa, Beyeler, Walt, & Barton, Dianne. (2004). Assessing infrastructure interdependencies: The challenge of risk analysis for complex adaptive systems. *International Journal Of Critical Infrastructures, 1*(1), Pp108-117.

Carey, Gemma, Malbon, Eleanor, Carey, Nicole, Joyce, Andrew, Crammond, Brad, & Carey, Alan. (2015). Systems science and systems thinking for public health: A systematic review of the field. *BMJ Open, 5*(12), E009002.

Checkland, P. (1999). *Systems thinking, systems practice*. Chichester: Wiley.

Chifari, Rosaria, Lo Piano, Samuele, Bukkens, Sandra G.F, & Giampietro, Mario. (2018). A holistic framework for the integrated assessment of urban waste management systems. *Ecological Indicators, 94*, 24-36.

Clayton, A. M. H., & Radcliffe, N. J. (2015). *Sustainability: a systems approach*. London: Routledge.

Conn, Katharine M. (2017). Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Impact Evaluations. *Review of Educational Research, 87*(5), 863-898.

Crawford, Lee. 2017. School Management and Public-Private Partnerships in Uganda. RISE Working Paper 17/013. www.riseprogramme.org/publications/rise-working-paper-17013-school-management-and-public-private-partnerships-uganda.

D'Agostino, G., & Scala, A. 2014. *Networks of Networks: The Last Frontier of Complexity*. Cham: Springer International Publishing: Imprint: Springer.

De Ree, J, 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.

De Savigny, D., & Adam, T. 2009. *Systems thinking for health systems strengthening* (Alliance Flagship report series). Geneva: World Health Organization.

Dudenhoeffer, D. D., Permann, M. R., & Manic, M. 2006. *CIMS: A framework for infrastructure interdependency modelling and analysis.*

Eddy-Spicer D, Ehren M, Bangpan M, Khatwa M, Perrone F. (2016). Under what conditions do inspection, monitoring and assessment improve system efficiency, service delivery and learning outcomes for the poorest and most marginalised? A realist synthesis of school accountability in low- and middle-income countries. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Eusgeld I., & Kröger W. (2008). Towards a framework for vulnerability analysis of interconnected infrastructures. In: Proceedings of the ninth international conference on probabilistic safety assessment and management (PSAM 9). p. 107–16.

Evans, P. B. (1995). *Embedded autonomy: states and industrial transformation*. Princeton, N.J.: Princeton University Press.

Frenk, Julio. (1994). Dimensions of health system reform. *Health Policy (Amsterdam), 27*(1), 19-34.

Frenk, Julio. (2010). The global health system: Strengthening national health systems as the next step for global progress. *PLoS Medicine, 7*(1), E1000089.

Fulop N, Allen P, Clarke A, Black N (eds). 2001. *Studying the Organisation and Delivery of Health Services, Research methods*. London: Routledge Publishers.

Forrester, Jay W. (1961). *Industrial dynamics*. Cambridge, Mass. : New York ; London: MIT Press ; Wiley.

Forrester, Jay W. (1968). *Principles of Systems.* (2nd ed.). Portland, OR: Productivity Press.

Gilson, Lucy. (2003). Trust and the development of health care as a social institution. *Social Science & Medicine (1982), 56*(7), 1453-1468.

Gilson, Lucy Ed. 2012. Health Policy and Systems Research - A Methodology Reader. *World Health Organization.* www.who.int/alliance-hpsr/resources/reader/en/.

Gilson, Lucy, & Raphaely, Nika. (2008). The terrain of health policy analysis in low- and middle-income countries: A review of published literature 1994–2007. *Health Policy and Planning, 23*(5), 294-307.

Gordon, T. J., & Helmer, O. 1964. *Report on a Long-Range Forecasting Study.* Santa Monica, CA: RAND P-2982.

Greenhalgh, Trisha, Macfarlane, Fraser, Steed, Liz, & Walton, Robert. (2016). What works for whom in pharmacist-led smoking cessation support: Realist review. *BMC Medicine, 14*(1), 209.

Greenhalgh, Trisha, Wherton, Joseph, Papoutsi, Chrysanthi, Lynch, Jennifer, Hughes, Gemma, A'Court, Christine, . . . Shaw, Sara. (2017). Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. *Journal of Medical Internet Research, 19(11), E367.*

George, Asha. (2009). 'By papers and pens, you can only do so much': Views about accountability and human resource management from Indian government health administrators and workers. *The International Journal of Health Planning and Management,24*(3), 205-224.

Glewwe, Paul, and Karthik Muralidharan. 2015. *Improving school education outcomes in developing countries: evidence, knowledge gaps, and policy implications*. University of Oxford, Research on Improving Systems of Education (RISE)

Haddad, E. A., Perobelli, F. S., Domingues, E. P., & Aguiar, M. 2011. *Assessing the ex-ante economic impacts of transportation infrastructure policies in Brazil*. Journal of Development Effectiveness, 3(1), 44–61. https://doi.org/10.1080/19439342.2010.545891

Hall, A. (1962). *A methodology for systems engineering*. Princeton: Van Nostrand.

Hall, Jim W., Scott Thacker, Matt C. Ives, Yue Cao, Modassar Chaudry, Simon P. Blainey, and Edward J. Oughton. (2017). "Strategic Analysis of the Future of National Infrastructure." *Proceedings of the Institution of Civil Engineers*: 1-9.

Hallegatte, S., Shah, A., Lempert, R., Brown, C., & Gill, S. 2012. *Investment Decision Making Under Deep Uncertainty: Application to Climate Change*. Policy Research Working Paper, (6193), 41. http://doi.org/doi:10.1596/1813-9450-6193

Halsey, Rogers, & Demas, Angela. 2013. The What, Why, and How of the Systems Approach for Better Education Results (SABER). *The World Bank*, 2 Aug. 2013, documents.worldbank.org/curated/en/2013/04/18070354/systems-approach-better-education-results-saber.

Hammerschmid, Gerhard *et al*. 2016. *Public Sector Reforms in Europe: The View from the Top.*

Hanson, Kara. (2015). What Can Education Systems Research Learn from Health Systems Research? RISE Working Paper 15/003. www.gov.uk/dfid-research-outputs/rise-working-paper-15-003-what-can-education-systems-research-learn-from-health-systems-research.

Hansen, H., Andersen, O. W., & White, H. 2011. *Impact evaluation of infrastructure interventions*. Journal of Development Effectiveness. http://doi.org/10.1080/19439342.2011.547659

Harris, G. (2010). *Seeking Sustainability in an Age of Complexity*. Journal of Environmental Management and Tourism, 1(1), 63–78. http://doi.org/10.1017/CBO9780511815140

Hawe, P., Shiell, A., & Riley, T. 2004. *Complex interventions: how "out of control" can a randomised controlled trial be?* BMJ, 328(7455), 1561. https://doi.org/10.1136/bmj.328.7455.1561

Hawe, P. 2015. *Lessons from Complex Interventions to Improve Health*. Annual Review of Public Health, 36, 307-323.

Heller, P. S. 1982. *A model of the demand for medical and health services in Peninsular Malaysia*. Social Science & Medicine, 16(3), 267–284. https://doi.org/10.1016/0277-9536(82)90337-9

Hirsch, G., Homer, J., Evans, E., & Zielinski, A. 2010. *A system dynamics model for planning cardiovascular disease interventions*. American Journal of Public Health, 100(4), 616. https://doi.org/10.2105/AJPH.2009.159434

Homer, Jack B, & Hirsch, Gary B. (2006). System Dynamics Modeling for Public Health: Background and Opportunities. *American Journal of Public Health (1971), 96*(3), 452-458.

Foundations for Decision Making. In *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 195–228). http://doi.org/10.1017/CBO9781107415379.007

Jackson, M C. (2009). Fifty years of systems thinking for management. *The Journal of the Operational Research Society, 60*(S1), S24-S32.

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

Kalra, N., Hallegatte, S., Lempert, R., Brown, C., Fozzard, A., Gill, S., & Shah, A. (2015). Agreeing on Robust Decisions New Processes for Decision Making Under Deep Uncertainty. *World Bank Policy Research Working Paper*, *No. 6906*(June). http://doi.org/doi:10.1596/1813-9450-6906

Kremer, Michael, & Holla, Alaka. (2009). Improving education in the developing world. *Annual Review of Economics, 1*(1), 513-542.

Law, Eugene P, Diemont, Stewart A.W, & Toland, Timothy R. (2017). A sustainability comparison of green infrastructure interventions using emergy evaluation. *Journal of Cleaner Production, 145*, 374-385.

Lempert, Robert J, Groves, David G, Popper, Steven W, & Bankes, Steve C. (2006). A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios. *Management Science, 52*(4), 514-528.

Lempert, Robert & Kalra, Nidhi & Peyraud, Suzanne & Mao, Zhimin & Tan, Sinh & Cira, Dean & Lotsch, Alexander. (2013). Ensuring Robust Flood Risk Management in Ho Chi Minh City. World Bank.

Lempert, R., Popper, S., & Bankes, S. (2003). *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis and bibliography* (Ebook central). Santa Monica, CA: RAND.

Lempert, Robert, Scheffran, Jürgen, & Sprinz, Detlef F. (2009). Methods for Long-Term Environmental Policy Challenges. *Global Environmental Politics, 9*(3), 106-133.

Levy, Brian, Cameron, Robert, Hoadley, Ursula, & Naidoo, Vinothan. (2018). *The Politics and Governance of Basic Education*. Oxford: Oxford University Press.

Lewis, T. (2006). *Critical infrastructure protection in homeland security: Defending a networked nation*. Hoboken, N.J.: Wiley-Interscience.

Lich, Kristen Hassmiller, Ginexi, Elizabeth M, Osgood, Nathaniel D, & Mabry, Patricia L. (2013). A Call to Address Complexity in Prevention Science Research. *Prevention Science, 14*(3), 279-289.

Mackenzie, M., Koshy, P., Leslie, W., Lean, M., & Hankey, C. 2009. Getting beyond outcomes: A realist approach to help understand the impact of a nutritional intervention during smoking cessation. *European Journal of Clinical Nutrition*, 63, 1136-1142.

Magrath, B., Aslam, M., & Johnson, D. (2019). Systems Research in Education: Designs and methods. *Research in Comparative and International Education, 14*(1), 7-29.

Marchal, Bruno, Van Belle, Sara, Van Olmen, Josefien, Hoerée, Tom, & Kegels, Guy. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation (London, England. 1995),18*(2), 192-212.

Marchal, Bruno, Westhorp, Gill, Wong, Geoff, Van Belle, Sara, Greenhalgh, Trisha, Kegels, Guy, & Pawson, Ray. (2013). Realist RCTs of complex interventions – An oxymoron. *Social Science & Medicine (1982), 94*, 124-128.

Masood, Tariq, McFarlane, Duncan, Parlikad, Ajith Kumar, Dora, John, Ellis, Andrew, & Schooling, Jennifer. (2016). Towards the future-proofing of UK infrastructure. *Infrastructure Asset Management, 3*(1), Pp28-41.

Mills, Anne. (2012). Health policy and systems research: Defining the terrain; identifying the methods. *Health Policy and Planning, 27*(1), 1-7.

Masset, E. (2019). Impossible generalizations: meta-analyses of education interventions in international development. RISE Programme.
https://riseprogramme.org/sites/default/files/inline-files/Masset%2011052019.pdf

Medical Research Council. 2000. *A framework for the development and evaluation of randomised controlled trials for complex interventions to improve health*. London: MRC, 2000.

McDaniel, Jr, Reuben R, Lanham, Holly Jordan, & Anderson, Ruth A. (2009). Implications of complex adaptive systems theory for the design of research on health care organizations. *Health Care Management Review, 34*(2), 191-199.

Moore, Mark. 2015. Creating Efficient, Effective, and Just Educational Systems through Multi-Sector Strategies of Reform. RISE Working Paper 15/004. https://riseprogramme.org/sites/default/files/2020-11/RISE_WP-004_Moore-REV%20copy.pdf

Ottens, Maarten, Franssen, Maarten, Kroes, Peter, & Van De Poel, Ibo. (2006). Modelling infrastructures as socio-technical systems. *International Journal Of Critical Infrastructures, 2*(2/3), Pp133-145.

Otto, Alexander, Hall, Jim W, Hickford, Adrian J, Nicholls, Robert J, Alderson, David, Barr, Stuart, & Tran, Martino. (2016). A Quantified System-of-Systems Modeling Framework for Robust National Infrastructure Planning. *IEEE Systems Journal, 10*(2), 385-396.

Ouyang, Min. (2014). Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering & System Safety, 121*, 43-60.

Pawson, R., & Tilley, N. 1997. *Realistic evaluation*. London; Thousand Oaks, Calif.: Sage.

Pawson R, Greenhalgh T, Harvey G, Walshe K (2005) Realist review: A new method of systematic review designed for complex policy interventions. *Journal of Health Services Research and Policy* 10: 21-34.

Pritchett, Lant, & Sandefur, Justin. (2015). Learning from Experiments when Context Matters. *The American Economic Review, 105*(5), 471-475.

Pritchett, Lant. 2015a. Creating Education Systems Coherent for Learning Outcomes. *RISE Working Paper 15/005* . www.riseprogramme.org/publications/rise-working-paper-15005-creating-education-systems-coherent-learning-outcomes.

Pritchett, Lant. 2015b. RISE: Research on Improving Systems of Education – Information Meeting for Bidders. YouTube, 9 July 2015, www.youtube.com/watch?v=dEQTUDhtMXg.

Pritchett, Lant. 2018. What We Learned from Our RISE Baseline Diagnostic Exercise. Rise Programme. www.riseprogramme.org/blog/baseline_diagnostic_exercise_1.

Rinaldi, S.M, Peerenboom, J.P, & Kelly, T.K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems, 21*(6), 11-25.

Rogers, Patricia J. (2000). Causal models in program theory evaluation. *New Directions for Evaluation, 2000*(87), 47-55.

Rogers, Patricia J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation (London, England. 1995), 14*(1), 29-48.

Rossiter, Jack., Woodhead, Martin., Rolleston, Caine., & Moore, Rhiannon. (2018). Delivering on every child's right to basic skills, Summative Report. Oxford: Young Lives.

Ryan, Mandy. (2004). Discrete choice experiments in health care. *BMJ, 328*(7436), 360-361

Saidi, Saeid, Kattan, Lina, Jayasinghe, Poornima, Hettiaratchi, Patrick, & Taron, Joshua. (2018). Integrated infrastructure systems—A review. *Sustainable Cities and Society, 36*, 1-11.

Seadon, Jeffrey K. (2010). Sustainable waste management systems. *Journal of Cleaner Production, 18*(16), 1639-1651.

Sheikh, Kabir, Gilson, Lucy, Agyepong, Irene Akua, Hanson, Kara, Ssengooba, Freddie, & Bennett, Sara. (2011). Building the field of health policy and systems research: Framing the questions. *PLoS Medicine, 8*(8), E1001073.

Sheikh, Kabir, & Porter, John. (2010). Discursive gaps in the implementation of public health policy guidelines in India: The case of HIV testing. *Social Science & Medicine (1982), 71*(11), 2005-2013.

Snyder, Sean. (2013). The Simple, the Complicated, and the Complex: Educational Reform Through the Lens of Complexity Theory. *OECD Education Working Papers*, No. 96, OECD Publishing. http://dx.doi.org/10.1787/5k3txnpt1lnr-en

Sturmberg, Joachim P, & Martin, Carmel M. (2009). Complexity and health - yesterday's traditions, tomorrow's future. *Journal of Evaluation in Clinical Practice, 15*(3), 543-548.

Thacker, Scott, Pant, Raghav, & Hall, Jim W. (2017). System-of-systems formulation and disruption analysis for multi-scale critical national infrastructures. *Reliability Engineering & System Safety, 167*, 30-41.

Thurlings, Marieke, & Brok, den, Perry. (2018). Student teachers' and in-service teachers' peer learning: A realist synthesis. *Educational Research and Evaluation, 24*(1-2), 13-50.

Tsofa, Benjamin, Molyneux, Sassy, Gilson, Lucy, & Goodman, Catherine. (2017). How does decentralisation affect health sector planning and financial management? a case study of early effects of devolution in Kilifi County, Kenya. *International Journal for Equity in Health,16*(1), 151.

Travis, Phyllida, Bennett, Sara, Haines, Andy, Pang, Tikki, Bhutta, Zulfiqar, Hyder, Adnan A, . . . Evans, Timothy. (2004). Overcoming health-systems constraints to achieve the Millennium Development Goals. *The Lancet (British Edition), 364*(9437), 900-906.

Tuominen, Anu, Tapio, Petri, Varho, Vilja, Järvi, Tuuli, & Banister, David. (2014). Pluralistic backcasting: Integrating multiple visions with policy packages for transport climate policy. *Futures : The Journal of Policy, Planning and Futures Studies, 60*, 41-58.

Von Bertalanffy, L. (1972). The History and Status of General Systems Theory. *Academy of Management Journal, 15*(4), 407-426.

World Bank. 2014. *SABER in Action: An Overview - Strengthening Education Systems to Achieve Learning for All*.
documents.worldbank.org/curated/en/866881468323335358/SABER-in-Action-An-Overview-Strengthening-Education-Systems-to-Achieve-Learning-for-All.

Walker, S., Gin, W., & Branch, E. S. 2012. *CalRecycle Review of Waste-to-Energy and Avoided Landfill Methane Emissions*. Retrieved from http://dpw.lacounty.gov/epd/conversiontechnology/download/CalRecycle_Review_of_WtE_Avoided_Emissions_07032012.pdf

Wilkinson, A., & Eidinow, E. 2008. *Evolving practices in environmental scenarios: A new scenario typology*. Environmental Research Letters. http://doi.org/10.1088/1748-9326/3/4/045017

Woo-Cumings, M. (1999). *The developmental state*. London: Cornell University Press.

Wong, Geoff, Greenhalgh, Trish, Westhorp, Gill, Buckingham, Jeanette, & Pawson, Ray. (2013). RAMESES publication standards: Realist syntheses. *BMC Medicine, 11*(1), 21.

World Bank. (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Washington, DC: World Bank. https://doi.org/10.1596/978-1-4648-1096-1

World Health Organization. 2002. *Strengthening Health Systems: The Role and Promise of Health Policy and Systems Research*. www.who.int/alliance-hpsr/resources/publications/hssfr/en/.

Woldehanna, Tessew, and Mesele W. Araya. 2017. *Early Investment in Preschool and Completion of Secondary Education in Ethiopia: Lessons From Young Lives*. www.younglives.org.uk, 8 Mar. 2017, www.younglives.org.uk/node/8415.

Xiao, Nan & Sharman, Raj & Rao, Raghav & Upadhyaya, Shambhu. (2008). Infrastructure Interdependencies Modeling and Analysis - A Review and Synthesis. 14th Americas Conference on Information Systems, AMCIS 2008. 1. 224.

Zhang, Pengcheng, & Peeta, Srinivas. (2011). A generalized modeling framework to analyze interdependencies among infrastructure systems. *Transportation Research. Part B: Methodological, 45*(3), 553-579.