



RESEARCH ON IMPROVING
SYSTEMS OF EDUCATION

WORKING PAPER
October 2015

What can Education Systems Research Learn from Health Systems Research?

Kara Hanson

RISE-WP-15/003
28 October 2015



Funded by:



The findings, interpretations, and conclusions expressed in RISE Working Papers are entirely those of the author(s). Copyright for RISE Working Papers remains with the author(s). www.rise.ox.ac.uk

What Can Education Systems Research Learn from Health Systems Research?

Kara Hanson

London School of Hygiene and Tropical Medicine

Abstract: Effective delivery of health services is underpinned by a health system, and the field of health systems research has emerged over the past 15-20 years as a recognised field of academic enquiry. The aim of this paper is to feed into the first year of conceptual thinking by the Intellectual Leadership Team of the Research for Improving Systems of Education (RISE) programme, which seeks to shape a research agenda to strengthen education systems in low- and middle-income countries. The paper reviews the development of the health systems research field, exploring the issues of system boundaries, alternative health system models, the range of disciplines and methods required for systems research, the contributions of health systems research to policy development, and the structures which could be developed to support a new field of education systems research.

Introduction

Much of health research focuses on development of specific technologies and interventions. Many of these interventions have demonstrated their potential to save lives, and have been shown to be cost-effective and therefore “good buys” for national governments and their development partners. Yet it is increasingly recognized that these interventions are delivered through a system, which in many settings is weak. Population coverage of many interventions is inadequate, quality is poor and may not meet either technical standards or patient expectations, resources are allocated to less effective and more costly activities, and households often pay substantial sums out-of-pocket to secure needed services, exposing them to risk of impoverishment. Weak health systems are highly vulnerable to shocks, as the recent Ebola epidemic in West Africa demonstrates.

Despite the overwhelming importance of health systems, the tendency to focus research funding in new technologies and to neglect research on the systems which are responsible for delivering these technologies persists:

“A recent study estimated the potential returns to investment in new technology versus research that could improve the delivery and utilization of health services. Surveying deaths among children under five years of age, in 42 low-income countries, the authors concluded that while improved technology had the potential to avert 21.5% of potential deaths, improved service utilization could avert 62.5% of child deaths. Despite the much greater returns to research on service delivery and utilization, the same study found that 97% of the grants awarded by the two largest research funders were for the development of new technologies”.

- (Alliance for Health Policy and Systems Research 2007) citing (Leroy, JP et al. 2007)

And the Director-General of the World Health Organization has also recognised the need for greater emphasis on systems research:

“Public health [today] enjoys commitment, resources, and powerful interventions...but the power of these interventions is not matched by the power of health systems to deliver them to those in greatest need, on an adequate scale and on time...This arises, in part, from the fact that research on health systems has been so badly neglected and underfunded...In the absence of sound evidence, we will have no good way to compel efficient investments in health systems.”

- Dr Margaret Chan, Director-General, World Health Organization, Beijing, China, 29 October 2007, cited in (Alliance for Health Policy and Systems Research 2012).

Health systems research has developed over the last 15-20 years as a recognised field of academic enquiry aimed at addressing these gaps. It has its own journals (e.g. *Health Policy and Planning*, tagline “The journal of health policy and systems research”; and *Health Systems and Reform*, launched in January 2015), a biennial conference, the Symposium on Health Systems Research; and most recently, a professional society, Health Systems Global (<http://www.healthsystemsglobal.org/>). The Alliance for Health Policy and Systems Research, sitting within the World Health Organization, was created to promote the generation and use of health systems research to improve the health systems of low- and middle-income countries (www.who.int/alliance-hpsr/en/). While there is a range of definitions of the field and its focus, there is broad consensus that the aim of health systems research is to “generate new knowledge on the way that societies organise themselves to achieve health goals” (Alliance for Health Policy and Systems Research), and that the focus of health systems research is on understanding the cross-cutting “building blocks” of health systems, and the linkages among them, rather than on narrow questions about a particular disease or condition. Underlying all of these elements is a focus on the *systems* required to deliver improvements in health and financial protection, which makes the field of health systems research a potential source of learning for the emerging field of education systems research.

From the perspective of one unfamiliar with the details of how education systems operate, health and education systems appear to have some similarities: they are large, national public services that frequently carry universal commitments (a right to health, to primary education). These universal commitments together with the presence of positive externalities and other market failures mean that governments play a substantial role in both sectors. In both, though to different degrees, the private sector also plays a significant role shaped by the perceived gaps in public sector provision. There are challenges common to both sectors: ensuring that human resources (teachers, principals, nurses, doctors) are available throughout the country including in remote rural areas, and that these workers receive needed support and supervision to deliver a quality service; managing out-of-pocket costs to households (whether as formal fees or as other charges) so that they do not act as a deterrent to access and do not expose households to excessive financial burden; designing governance structures that both ensure accountability to communities and permit meaningful opportunities for citizens to express their views and preferences; ensuring adequate supplies (drugs, textbooks) in all parts of the country; and ensuring that quality standards are met across the sector, by both public and private providers.

There are clearly some differences between the sectors, too, such that different areas of research will receive different emphasis (Colclough C 1997). Education is cumulative while most health care is designed to restore health to a previous state; education can be provided in a fairly uniform way to groups of students while curative care, at least, is individually tailored; and the underlying demand for services is influenced by different factors, because of the intrinsically unpredictable nature of the demand for health care from which the demand for insurance is derived (de Moura Castro C and Musgrove 2002).

The performance of both health and education systems in low- and middle-income countries is mixed. PISA and TIMSS data show considerable shortfalls in achievement in mathematics compared to those in high-income countries (Pritchett 2013). In health, maternal mortality, widely accepted as a good indicator of health system performance because most maternal mortality is avoidable through health care intervention, is many times higher in low- and middle-income countries (LMICs)

than in high-income settings. Yet, at least in health, there is some variation in achievement for a given level of national income, suggesting that even with limited resources it is possible to do better (Balabanova D, Mills A et al. 2013). The specific way in which the health system is organised and financed has been shown to influence system-level performance (OECD 2004), and is amenable to change regardless of the absolute level of resources available.

Generating new knowledge to improve the performance of these systems should be the overall goal of health and education systems research. The overall aim of this paper is to feed into the first year of conceptual thinking by the RISE Intellectual Leadership Team as it seeks to frame the RISE research programme. It addresses issues of boundaries of the system, the range of disciplines and methods required, and the potential structures to support a field of education systems research. It does so by means of encouraging comparison with the (more established) field of health systems research.

The structure of the paper is as follows. Section 1 traces the origins of the field of health systems research, and identifies its conceptual boundaries. Section 2 provides an overview of some of the most commonly used models of health systems, concluding with an overview of health systems as complex adaptive systems and the research methods that are being applied to understand better this complexity. Section 3 considers how overall health system performance has been assessed, and some of the limitations of these approaches. Section 4 identifies a number of system features that are common to health and education. Section 5 focuses on the implications of intervention complexity for evaluation methods in systems research. Section 6 outlines some of the key challenges in health systems research and lessons for the emerging field of education systems research.

1. Health Systems Research – Origins and boundaries

As recently as 2008, a Google search of “health systems research” came up with no references other than to the Alliance website (Mills 2011), suggesting that it has only lately been recognised as a distinct field of research. However, historian Martin Gorsky has traced an early use of the term “health system” to 1896 (in reference to the potential for adoption of models of “practical sanitation” by US local governments). Proposals for the creation of organised systems of medicine in the inter-war years variously used the term “health service” and “health system” though in common held a view of the need for (and value of) coordination by the state. The 1960s saw accelerated use of the term, as European systems of Universal Coverage were consolidated and the field of comparative health systems research developed (Gorsky 2013). In low- and middle-income countries, the idea of researching systems may have emerged from the shift from research of the post-colonial period, addressed at combatting major causes of illness through disease-specific intervention programmes, to a focus on primary care, service integration and comprehensive community-based services. While not using the term “health system”, the elements of primary health care included health education, provision of water and sanitation, nutrition and food supply, comprehensive preventive and curative services, and provision of essential drugs; and research on how to strengthen primary health care included, for instance, a focus on cross-cutting elements such as community participation, all foreshadowing the cross-sectoral, cross-cutting nature of health system research.

In the 1980s the World Bank emerged as a key actor undertaking research and providing technical advice on health financing as one of the key building blocks of the health system. Notable

contributions at this time included the application of benefit incidence analysis to health, demonstrating the tendency for health system resources to be used disproportionately by better off groups (Castro-Leal, Dayton et al. 2002); and the publication in 1987 of the World Bank's "Agenda for Reform" which advocated a much stronger role for user fees to finance health services and manage "frivolous" demand for services, spurring a significant body of work on the impact of such charges on utilization (Gilson L 1988, Waddington C and Enyimaew 1990). The 1993 World Development Report, Investing in Health, was another key publication drawing attention to the underpinning health system functions, advocating the use of cost effectiveness analysis and estimates of the Global Burden of Disease to select an essential package of services to be delivered at different levels of the health system. In 2000, the World Health Report developed a framework to compare the performance of the health systems of the 192 WHO member states measuring the extent to which they delivered three key elements – improved health outcomes, financial risk protection, and "responsiveness" or patient satisfaction (see below).

The Millennium Development Goals, with their focus on reducing the burden of specific diseases, produced a resurgence of disease-specific funding streams such as the Global Fund for AIDS, Malaria and Tuberculosis (2002), major US disease-specific initiatives (The President's Emergency Programme for AIDS Relief (PEPFAR) in 2003, the President's Malaria Initiative in 2005), and GAVI Alliance, supporting vaccination programmes. But this was quickly followed by a recognition that disease-specific targets could not be met and sustained without a strong health system foundation (Travis 2004) and that a systems approach was needed to avoid the negative effects of competing and conflicting separately organized disease control programmes. The WHO Task Force for Health Systems Research, launched in 2003, noted that "Health system constraints constitute major barriers to achieving the MDGs; there are many unanswered questions about how to strengthen health systems and scale-up effective interventions; and more resources are needed to answer these questions and build capacity within less-developed countries" (http://www.who.int/rpc/summit/en/Task_Force_on_Health_Systems_Research.pdf).

The Alliance for Health Policy and Systems Research was the first organisation to add the term "policy" to describe the field, thereby explicitly recognising a role for political science and policy analysis in informing policy development. Policy analysis methods including stakeholder analysis have been used in health systems research to trace the patterns of support and opposition to reforms (e.g. In South Africa (Gilson, Bowa et al. 2000); and an early "how to" article on stakeholder analysis remains one of the most downloaded papers from *Health Policy and Planning* (Varvasovszky Z and Brugha R 2000).

This research *on* policy development has been accompanied by research *for* policy development, with the field strongly driven by a desire to shape policy. This involves engaging researchers who are interested in the real world, involving health sector actors (policymakers, managers, providers) in the research process, and being "policy-minded" in research outputs, reflecting on how results can be framed and located in the specific context and in actionable terms (Sheikh K, George et al. 2014). It has drawn on, and contributed to, the literature on "getting research into policy and practice" reflected in, for instance, an Alliance for Health Policy and Systems Research strategy document which focuses on how to conduct research so as to influence policy; and on what sorts of capacities needed by researchers and policymakers (Alliance for Health Policy and Systems Research 2012).

Box 1.

The Alliance strategy has developed four case studies (China, India, Mexico, Thailand) of countries where the use of HSPR in policymaking has been institutionalised. Some of the key lessons include:

- *Embed research within decisionmaking processes:* establish institutional mechanisms such as protocols for policy formulation, planning and implementation that explicitly refer to evidence. In Mexico, for example, legislation requires all large-scale social policies to be independently evaluated. This led to, for example, a rigorous approach to evaluating the Progresa / Oportunidades cash transfer programme which generated the evidence which helped to maintain it. Create opportunities for closer interaction between policymakers and researchers by, e.g., placing data in the public domain.
- *Support demand-driven research:* Elicit research priorities from a variety of stakeholders through transparent, systematic, and inclusive processes. In Thailand, the annual updates to the health insurance benefit package are undertaken through a process which includes consultation with patient and industry groups to identify new interventions and technologies which are then subject to economic evaluation to reach a decision about whether they are cost-effective and affordable.
- *Strengthen capacity for research and use of evidence:* This requires a long-term strategy of developing training programmes for researchers, developing methods, and supporting links between researchers to build the field. In addition there is a need to expose decisionmakers to the use of evidence in policymaking, through formal training and exposing them to the research process, for example, by rotating government staff through research institutions
- *Establish repositories of knowledge:* Knowledge synthesis is needed to make usable and accessible the vast body of research evidence generated, and to encourage sharing of evidence across countries.

Factors facilitating the successful development of a culture of evidence-informed policymaking included:

- establishing health policy and systems research institutions with able leadership and competent staff;
- involvement and influence of prominent health researchers and health leaders in advocating and highlighting the role and promise of HPSR;
- increased allocation of resources towards health policy and systems research;
- a prominent role of press and public opinion in highlighting important issues related to health policy and health systems;
- positive contribution of international funders in strengthening the generation of evidence and its use locally;
- support of parliamentary standing committees in demanding policy-relevant knowledge;
- orchestrated efforts to embed research into decision-making; use of evaluation as a tool for learning, transparency and the scale-up of programmes.

Source: (Alliance for Health Policy and Systems Research 2012)

Other relevant extensions to boundaries concern the inclusion of global level policies and actors as an important influence on national level policies and systems; and the question of whether the focus is the health *care* system or whether it should also encompass other sectors which influence health, such as trade and environmental policy (Smith and Hanson 2012).

Perhaps most importantly, the field of HSR has developed to be defined by the types of problems that it addresses and the questions it asks, rather than by a particular disciplinary perspective. It seeks to produce evidence to improve the coverage, quality, efficiency and equity of health systems. It is concerned with research questions that are descriptive and explanatory as well as evaluative (Sheikh et al.) It therefore draws on a wide variety of epistemological traditions, and therefore a broad set of research methods and approaches.

For all of these boundaries there are clear parallels with the field of education systems research. First, while there is quite a lot of evidence generation about interventions to improve learning outcomes, such as the use of IT, teacher training, and interventions to reduce class size (see for example, (McKewan PJ 2014)), there seems to be less research on the systems needed to ensure that these effective interventions are financed and delivered, and at scale. Second, the debates about the boundaries of the system seem germane to education – learning can be improved by non-education interventions such as nutrition supplementation or potentially, interventions to improve accountability to communities for school performance. Third, there are likely to be some cross-cutting factors that affect the delivery of all public services, such as administrative reforms, decentralization, and the operation of the civil service. Finally, research is also needed on the process of policy development and implementation, to provide insights on how best to manage the roles of multiple stakeholders such as teachers unions and their responses to organisational change.

2. Models for understanding health systems – no single one will suffice

Those involved in promoting the new area of education systems research often ask whether there is “a model” that is used to understand health systems. The field of health systems research has produced a multiplicity of models. Some look at the system as a whole while others look at specific sub-systems such as health financing or governance. Some are designed to characterise and describe a system by listing the organisations or functions, adopting an “inventory” approach (Frenk J 1994) while others specify the relationships among the components and actors (“relational” models) that allow the underlying institutional features of these relationships to be explored and, potentially for hypothesising and testing. (Gilson L 2012) contains a detailed exploration and critique of the different health system models and proposes specific extensions.

Inventory approaches and “building blocks”

Perhaps the most commonly used inventory model is known as the WHO health system building blocks model (World Health Organization 2007), which identifies the key health system functions and links these through a black-box-like mechanism of intermediate outcomes to the overall health system goals of improving health, responsiveness, risk protection and efficiency. Such models are useful for describing and characterising, but have been found unsatisfactory for characterising the linkages among health system functions, or for understanding health system performance.

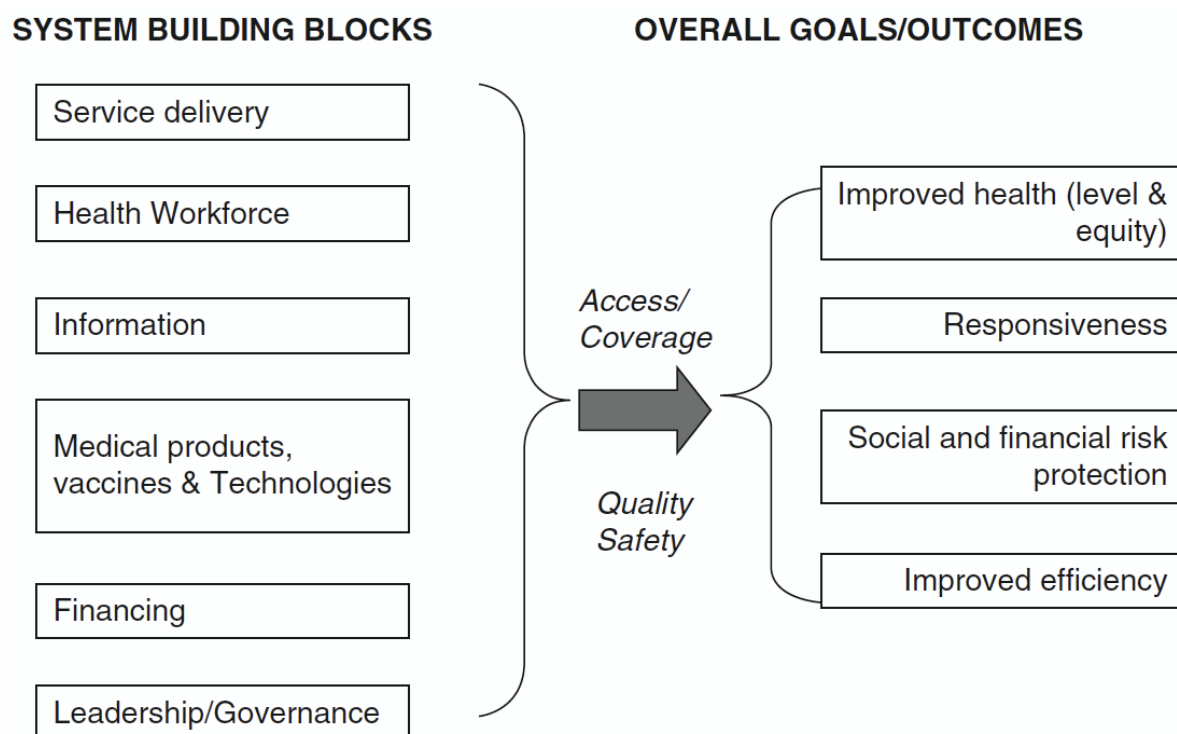


Figure 1 WHO Building block model - no permission sought

Building in relationships among actors/functions and the bases for influence

More detailed specification of the actors and relationships among them is possible when characterising a particular function/building block or sub-system, such as the health financing sub-system. WHO's model of the health financing system depicted in Figure 2 identifies the different actors involved (consumers, providers, purchasers, government and professional bodies) and the core financing functions (resource mobilization, resource allocation and purchasing, and service delivery; note that pooling of resources to allow cross-subsidy is not directly reflected here but is also a core function identified in other models). As (Gilson L 2012) highlights, this conception of the system also indirectly articulates the institutions underpinning the different relationships among actors – regulatory authority exercised by governments and professional bodies; payments by patients / population, reflecting economic incentives; and payments to providers, governed by rules, but also containing economic incentives.

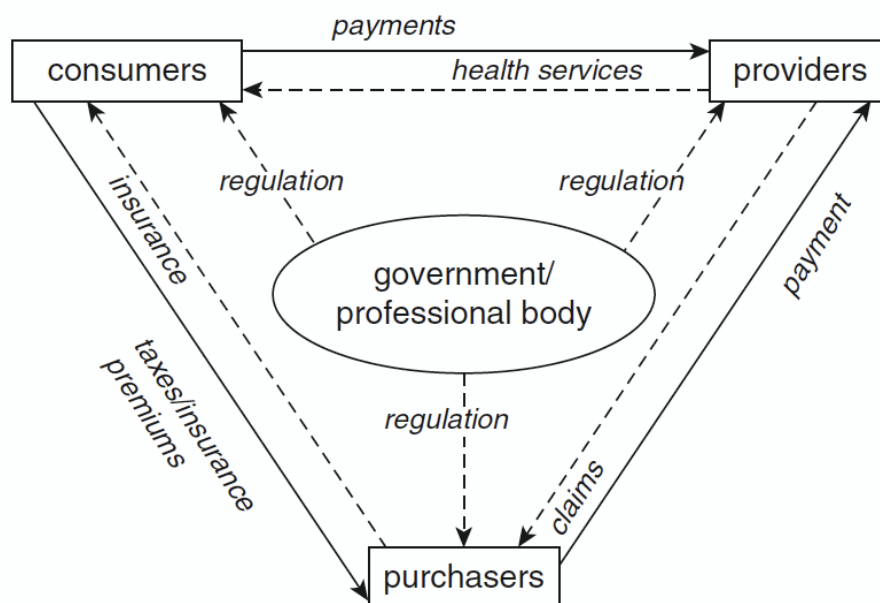


Figure 2 WHO Health Financing Model (no permission sought)

Frenk's (1994) health system model starts with the observation that all health systems are based on an interaction between provider and patient. However, this relationship is complicated and extended by a number of additional features. First, both providers and members of the population act through organisations which shape their interactions. In the case of patients, for example, these units of organisation can include households, communities, and political organisations and interest groups; for providers these might be professional associations or unions. Second, neither provider nor patient organisations are homogeneous: they reflect the various cleavages and stratifications of society at large, including, for instance, professional hierarchies. Third, the relationship between provider and patient is frequently mediated by a collective actor, usually the state, which can act as a regulator, a financing agent or a direct provider of services. The state also usually determines conditions of eligibility for the services that it controls. Finally, economic sectors other than the health sector influence health: these include sectors that traditionally lie within the "social determinants" of health such as food, water and sanitation, etc. But system-level research on health also increasingly needs to consider the effects of the broader policy environment, such as international trade in food and in health services, that influence health (Smith and Hanson 2011) (Figure 3).

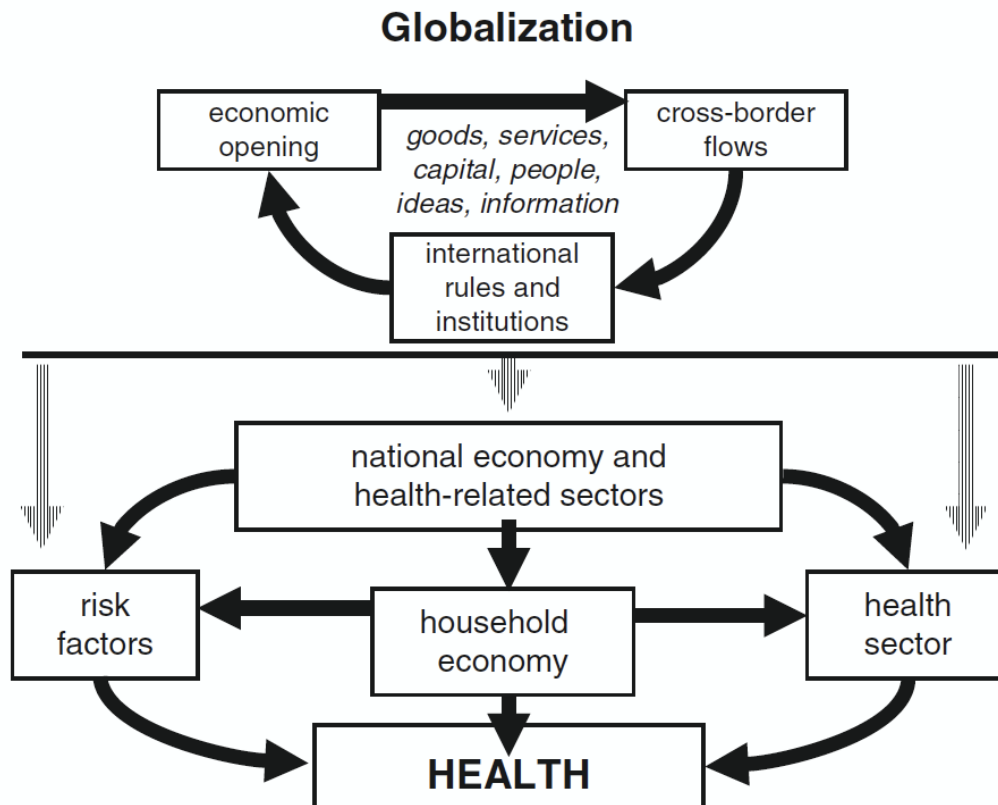


Figure 3 Smith and Hanson 2012; permission not sought

A slightly different approach is taken by (Roberts MJ, Hsiao et al. 2008) in their health system model, focusing on the “control knobs” or “power mechanisms” that influence how well the health system components function to achieve their intermediate performance measures and overall performance goals. Table 1, reproduced from Gilson 2012, illustrates how the control knobs map onto policy levers that can be exercised by governments to improve health system performance.

Table 1: Health system control knobs

Control knob	Influences
Financing	Who pays for and who benefits from healthcare, and the generation of funding for the system as a whole
Payment	How money is transferred to providers, creating financial incentives and influencing how they behave
Regulation	Use of the coercive power of the state to control the behaviour of actors within the system
Organisation	The incentives for the organisation; and incentives, authority, skills and attitudes of workers and managers
Behaviour	Information provision and marketing, incentives and coercion shaping how patients and providers act in relation to health and health care – influencing treatment seeking, health professionals behaviour, patient compliance, lifestyle and preventive behaviours

Reproduced from (Gilson L 2012). Permission not yet sought.

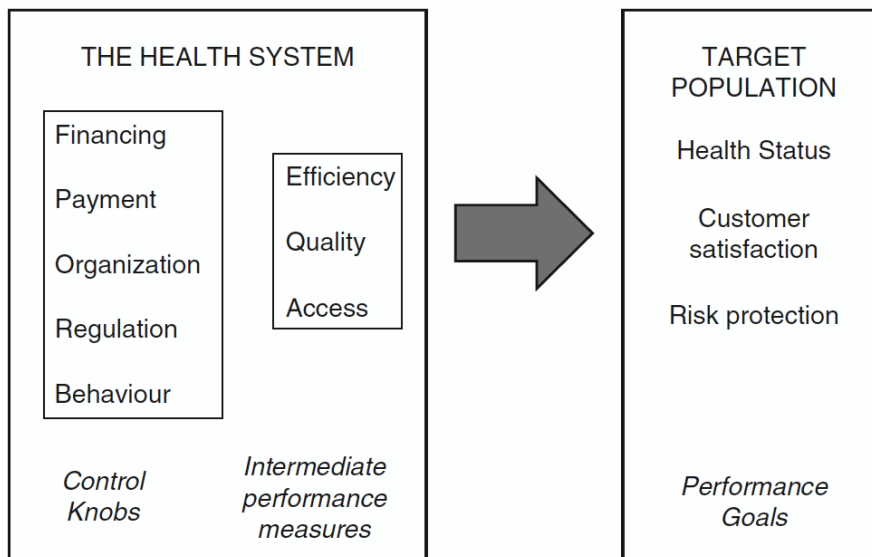


Figure 4 Roberts et al. "control knobs" model; no permission sought

Integrating systems "software"

Commenting on the development of health policy and systems research as a field, (Sheikh K, Gilson L et al. 2011) have emphasized the need to encompass the "software" of health systems as well as the hardware of systems and functions. Health systems software includes both the tangible software (human resources, management knowledge and skills, and formal management processes) and the intangible software (the informal rules, communication patterns, values and norms that shape relationships and interactions among actors)(Elloker S, Olckers P et al. 2013). Furthermore, by acknowledging that policies are implemented within a specific social and political context, the health systems model presented in the right-hand circle of Figure 5 recognises that interpretations of policy are socially constructed and therefore subject to multiple understandings. The application of theories of sensemaking in analysing the challenges of mid-level management in a South African district health system(Gilson L, Elloker S et al. 2014) and Street Level Bureaucracy to community-based distributors of contraceptives in Kenya (Kaler A and Watkins S 2001) provide insights into how gaps between central level health system initiatives and actual practice emerge as they are reinvented at local level.

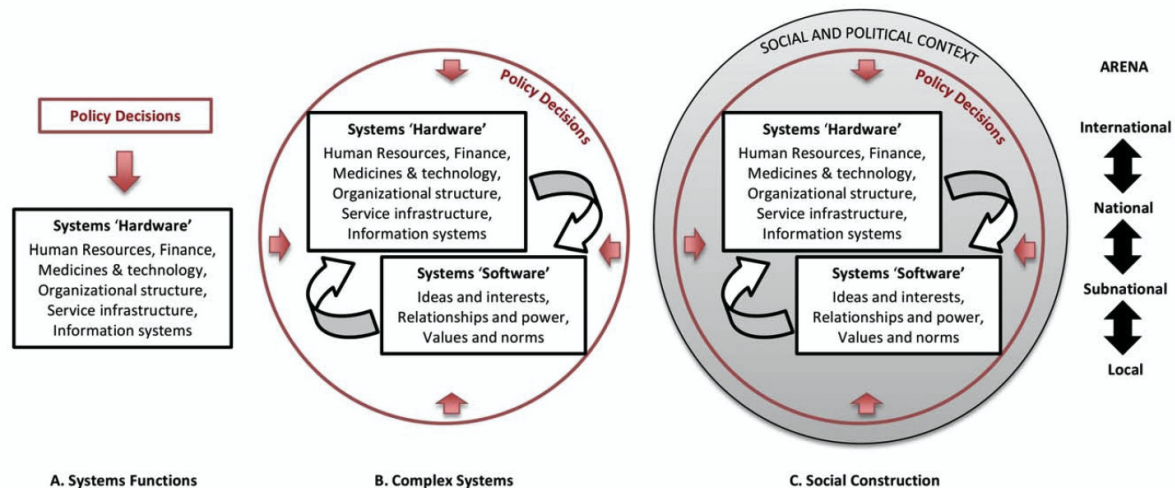


Figure 5 Sheikh et al. PLOS - integrating systems software and social construction of experience

In searching for appropriate models, the field of education systems research would do well to recognise that no single model will be able to frame all systems questions.

Health systems and complexity

Complexity theory is the most recent addition to the set of perspectives through which health systems are conceptualized and modelled. Better described as a set of approaches and models, the underlying idea is that health systems are complex and therefore require more flexible and adaptive models and research methods to accommodate their key properties, summarized in (Smith and Hanson 2011):

- They are made up of elements, individuals or organisations, connected together to form a whole and that itself possesses properties distinct from those of its component parts
- They are nested within other systems, so change in one sub-system affects other sub-systems, making it difficult to determine system boundaries.
- Systems are complex and adaptive: *‘a collection of individual agents with freedom to act in ways that are not always totally predictable, and whose actions are interconnected so that one agent’s actions changes the context for other agents’* (Plsek and Greenhalgh, 2001: 625).
- Feedback loops between agents, both direct and indirect, mean that agents influence other agents, but also that their own actions are fed back to them and this, in turn, affects the way they behave in the future. The past influences current behaviour, providing a basis for learning (Hudson, 2007).
- Agents also belong to several systems at the same time, and membership of such systems changes as agents move in and out of them. This means that change is constant in systems.

Path dependence is another commonly identified feature of complex systems, in which outcomes of interventions depend both on initial (current) conditions and on decisions taken in the past: policy choices and decisions about implementation modes are therefore influenced (and constrained) by earlier choices (Marchal B, Van Belle S et al. 2013). A 2014 special issue of Health Research Policy and Systems illustrates a variety of applications of systems thinking approaches to researching health systems. The research that is presented in these papers reflects a variety of analytical perspectives, ranging from descriptive and comparative through to evaluative; employs different theoretical perspectives, such as “sensemaking” and path dependency; and applies methods and

tools such as quantitative systems modelling to examine health resource allocation decisions (Bishai D, Paine L et al. 2014), social network analysis to look at the use of sustainability indicators for rehabilitation programmes in Nepal and Somiland (Blanchet K, Palmer J et al. 2014), and causal loop diagrams to explore the factors underlying changes in vaccine acceptability, including trust, in Kerala, India (Varghese J, Kutty V et al. 2014).

Despite this growing body of work, the implications of complexity science as an approach to health systems research are still being explored – both in terms of what types of research questions it can help answer, and what methods are best suited to different questions. The evidence from the health sector suggests that the methods of complexity science are well suited to explaining how health systems work in a nuanced and detailed way which can provide new insights. For example, complexity can be used to understand the constraints to improved health service performance as a precursor to a process of designing eventual interventions; for retrospective analyses of how and why reforms and policies are adopted; and for explaining why it is that so many well-meaning interventions do not have their anticipated effects, due to, for example, actor resistance and response.

However, the toolbox of complex systems analysis seems to be more limited in predicting the effects of future interventions. Peters (2014) argues that systems thinking is valuable to be able to explain how complex systems work; to provide guidance on where to collect more data, or to raise new questions and hypotheses; and to allow us to “understand and continuously test and revise our understanding of the nature of things, including how to intervene to improve people’s health” (p. 5). As discussed below, the methods of complexity theory can provide a critical addition to evaluation methods for complex interventions, but *on their own* seem unlikely to provide insights into the effectiveness of interventions delivered on a large scale.

More generally, even this partial review of the health system literature demonstrates that a highly varied set of models has been applied to health systems research questions. However, many of the models are descriptive, setting out an inventory of functions and actors rather than characterizing the relationships between them. Education system researchers will also need to develop a set of models adapted to the specific actors, functions and relationships in the education sector. But the field will also benefit from efforts to make such models more analytical, and particularly to focus on the policy signals (incentives, information flows, structures of accountability) that will stimulate and sustain system change.

3. Measuring system performance

As in the education sector with the TIMSS and PISA rankings, there have been some efforts to rank countries in terms of their health system performance with the aim of increasing accountability of governments and providing an evidence base to assist policymakers to improve system performance. Perhaps the most ambitious effort was the World Health Organization’s Health System Performance Assessment (HSPA), applied to all 192 member states and presented in the 2000 World Health Report (World Health Organization 2000). The conceptual model underlying the HSPA set out four main functions that health systems must carry out (financing, service provision, resource generation and stewardship or oversight of the system as a whole) (note that these are different from the WHO building blocks which followed). These functions together would determine the health system’s achievements in relation to three social goals: improving population health, improving responsiveness to population expectations, and fairness in financial contribution. The indicators chosen to capture attainment of these goals are shown in Table 2.

Table 2: Domains of WHO's Health System Performance Assessment

Domain	Indicator(s)
Health	Average health attainment of the population (disability-adjusted life expectancy, DALE) Distribution of health outcomes (child survival)
Responsiveness	Respect for persons (dignity, autonomy, confidentiality) Client orientation (prompt attention, quality of social amenities, access to social support networks during care, choice of provider All assessed through a survey of key informants; measured as level and distribution
Fairness in financial contribution	Fraction of disposable income contributed by households

Source: (World Health Organization 2000)

Country performance was reported as “goal attainment” – measured as performance on each of the 5 indicators individually and in a composite index; and “performance” which related achievements to expenditure, individually for health outcomes, and then comparing the composite index with what would be expected given the level of economic and educational performance. The league table which then compared country performance was highly controversial and prompted considerable debate. It is notable that WHO has not attempted to repeat or update the rankings in subsequent World Health reports; according to David Evans, a former WHO staff member who was involved in the development of the framework, the rankings and prompted much criticism of WHO and there was never any appetite within the leadership of the organization to repeat the process.

Data availability imposed a major constraint on the perceived validity of the various measures included in the HSPA – perhaps inevitable when attempting to rank the health system performance of all 192 member states. Lacking vital registration data, a considerable amount of modelling of health outcomes was needed, and health expenditure surveys to measure financial contributions were also scarce. In a critique of the methods of the report (Musgrove 2003) determined that only 39% of the indicators in the report were derived from existing sources; the remainder were imputed. The survey on which the responsiveness indicator was calculated was administered to only 1790 respondents in 35 countries, many of whom were WHO staff (Williams 2001).

Perhaps a greater limitation is the difficulty in relating the performance measures, irrespective of their limitations, to the organisation of the health system in order to understand what types of policies lead to higher levels of performance. Some analyses of specific dimensions of performance, particularly of health financing arrangements, have been possible. For instance, there has been analysis of the relationship between the share of healthcare payments that are out-of-pocket (indicating low levels of pooled expenditure, whether through insurance or tax funding) and the proportion of households that incur “catastrophic” payments (payments deemed large enough to send a household into, or more deeply into, poverty)(Xu, Evans et al. 2007). More recent analysis has demonstrated that systems that provide greater financial protection, measured as a smaller share of out-of-pocket payments in total health expenditure, have lower adult mortality(Moreno-Serra and Smith 2015). However, other features of health system organisation that are amenable to policy change, such as the role of private providers, the effectiveness of regulation, and the use of a primary care gatekeeper in enforcing the referral system, are not captured in a systematic manner

across countries, limiting the potential for such comparative assessments of system performance to generate meaningful policy lessons for countries.

The complexity of measures of health system performance, arising because of the multiple goals of a health system, contrast with the relative (!?) simplicity of measures to assess education system performance through the single outcome of student test scores. However, the number of countries participating in international comparisons of learning outcomes is much lower (70 economies for the OECD's PISA test; 63 countries in the 2011 TIMSS assessments). The results of these tests are used both for advocacy purposes to generate demand for system level change and for comparative analysis and lesson learning across countries (see for example (Woessmann L 2007, Pritchett and Viarengo M 2009, OECD Various)).

4. System-level interventions and reforms – parallels between health and education

A number of health system topics have parallels in education systems policy, and could form the basis for useful and informative comparative systems research.

Role of the private sector and how best to harness this for public goals: The private health sector is an important provider of health services, often chosen because it is more accessible, perceived to have higher quality and responsiveness, or be more affordable, than public sector providers. The sector is very diverse, however, ranging from international standard tertiary hospitals serving medical tourists through to untrained drug sellers, and there are concerns about the comprehensiveness of services offered, technical quality of care and ability of providers to induce demand for services given their higher powered incentives and information advantage. Private providers are very much part of the health system, though they frequently operate without much policy or regulatory oversight. Policy initiatives to harness the potential advantages of private provision have included strengthening private sector regulation, patient choice based initiatives including vouchers, and a host of interventions to improve quality of care such as training, clinical guidelines, franchising, accreditation, etc. There are also discussions about how to strengthen public sector purchasing to enable services to be procured from private providers, creating simultaneously a level playing field and a set of policy levers linked to payment that can be used to extract better performance from private providers. The private sector is also very active in education, from nursery level through to professional training institutions, and it seems equally to be responding to gaps that are perceived in the provision offered by the public sector. There is scope to update and extend the comparative analysis of the role of the private sector in education and health undertaken in (Colclough C 1997).

Improving attraction and retention of professionals in rural areas: Ensuring there are trained health workers in rural areas is an important policy to enhance equity. However, imbalances in the distribution of the workforce remain a challenge. Health systems researchers have conducted empirical research to understand the potential impact of policies to improve retention (Blaauw D, Erasmus E et al. 2010) and evidence has also been systematically reviewed and synthesised (World Health Organization 2010). These approaches could also be applied to policies to retain rural teaching staff. There is also scope to explore the feasibility of using longitudinal methods to study labour mobility, as has been done for a cohort of nursing school graduates in South Africa that has been followed now for more than 7 years (<http://resyst.lshtm.ac.uk/resources/infographic-tracking-nurses-job-choices-south-africa>).

Community accountability for performance and scorecards: Governance and accountability is one of the key health system building blocks and identifying ways to improve accountability of health providers to the communities in which they operate has been an important area of health system research (Molyneux CS, Atela M et al. 2012) and has been a focus also of education research. Improved accountability is argued to improve health worker responsiveness to local needs, and potentially to increase service utilization by prompting changes in service delivery. This is a further area where cross-sector collaboration could be fruitful.

5. Evaluating complex interventions – lessons from health

Health researchers have been at the forefront of developing methods for evaluating the impact of interventions – randomized controlled trials emerged from the field of medicine (the first RCT was published in 1940, of a TB treatment), and these experimental methods are now widely used across the social sciences. However, the complexity of health system interventions places new demands on evaluation methods, which health systems researchers are actively engaging with and extending. Three key areas stand out where health systems researchers have been drawing from the public health and broader evaluation literatures to advance the science of evaluating health system interventions: frameworks for evaluating “complex interventions”; the use of a theory of change; and realist evaluation.

As highlighted earlier, health system interventions frequently have several interacting components and involve change at multiple levels of the system. The UK Medical Research Council guidelines for evaluating complex interventions, recently updated with guidance for undertaking process evaluation (Moore GF, Audrey S et al. 2015), have been used to shape evaluations of health system interventions.

Table 3. Evaluation of complex interventions

Dimension of complexity	Implications for evaluation design
Number of and interactions between intervention components	A good theoretical understanding is needed of how the intervention causes change, so that weak links in the causal chain can be identified and strengthened
Number and difficulty of behaviours required by those delivering or receiving the intervention	Difficult to distinguish intervention ineffectiveness from implementation failure / teething problems: need process evaluation to identify implementation problems.
Number of groups or organisational levels targeted by the intervention	Variability in individual level outcomes may reflect higher-level processes; sample sizes may need to be larger to take account of the extra variability, and cluster- rather than individually-randomized designs considered.
Number and variability of outcomes	A single primary outcome may not be sufficient to understand impact; a range of measures will be needed, including unintended consequences identified from a theory of change
Degree of flexibility or tailoring of the intervention permitted	Ensuring strict fidelity to a protocol may be inappropriate; the intervention may work better if adaptation to local setting is allowed.

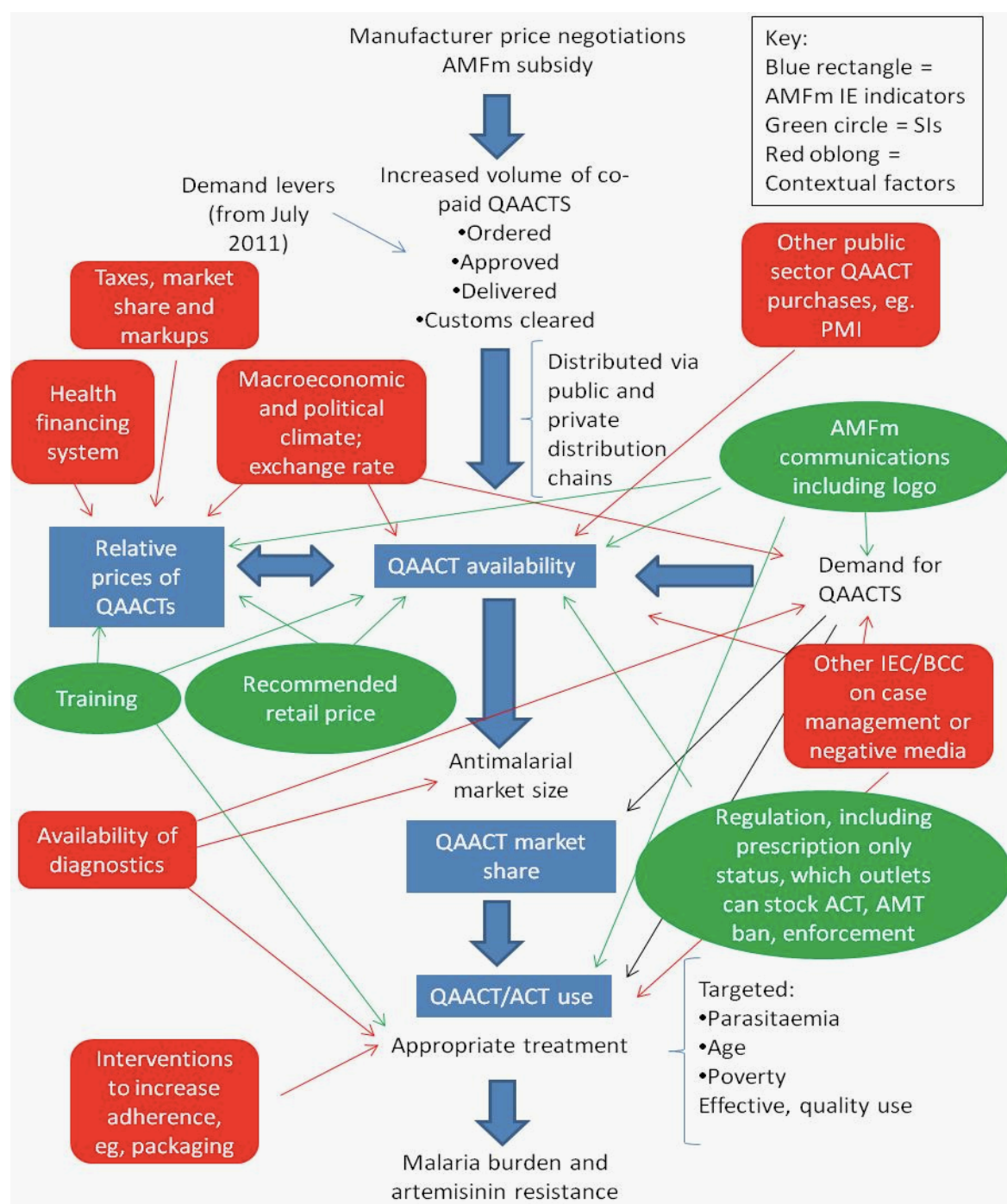
Adapted from MRC 2008.

Three of these characteristics stand out as adding to the conventional impact evaluation toolkit, which tends to focus on achieving an unbiased estimate of impact: These are the need for the evaluation to be guided by a clear theory of change; the use of process evaluation to support the testing of the theory of change, and to aid interpretation in the case of a negative result; and the need to include a range of outcomes, including unintended consequences. Together, these can help to understand how, why and for whom an intervention works.

Using a theory of change/programme theory for evaluation design and interpretation

Because of the complexity of health system interventions, evaluators are increasingly using an explicit theory of change to guide evaluation design and interpretation (see for example, Borghi et al. 2013, who present a theory of change to guide the evaluation of a pay-for-performance scheme in Tanzania). A theory of change, or programme theory, depicts how an intervention is supposed to work – in its most complete form it articulates the elements of the causal chain, together with any assumptions or preconditions that are required for the mechanism of action to operate. It may also specify potential unintended consequences and confounding factors that may be responsible for, or impede, change. For complex interventions a theory of change is useful for clarifying research questions and hypotheses, informing data requirements, and interpreting results, in other words, providing a structuring tool for an evaluation of a complex intervention (Marchal et al. 2013). By providing a systematic way of working through potential alternative explanations for change, it can also be helpful in informing judgements about attribution when the nature of the programme means that an experimental or quasi-experimental design is infeasible, contributing to the strength of plausibility statements (Habicht et al. 1999) by ruling out confounders.

Box 2: Using a theory of change to evaluate a malaria drug financing facility



The Affordable Medicines Facility –malaria (AMFm) was a health system intervention designed to reduce the price of effective antimalarial medicines through both the public and private sectors, thereby increasing their use. It was a complex intervention requiring actions at multiple levels of the system (supply by international pharmaceutical firms, orders from national level wholesalers, distribution through the private sector supply chain, stocking and sales by local retailers, appropriate demand and use by malaria patients or their caregivers). It also had multiple components: a subsidy applied at the top of the distribution chain, public information campaigns and provider training, and the use of a logo and recommended retail price to direct

demand towards the subsidised drugs and constrain mark-ups. The intervention and its evaluation were conducted in 8 national level pilot programmes, over two years.

A before-and-after study design was used, as it was not possible to identify an appropriate control group. In addition, a detailed assessment of the context was undertaken at baseline and endline, and a process evaluation captured indicators of implementation strength. Benchmarks for success were defined for each of the main outcomes (availability, price, market share and use). After implementation periods ranging from 5 to 15 months, benchmarks for price and availability had been met in 5 of the 8 pilots, for market share in 4 of the pilots, and for use in only one of the 5 in which use was measured.

In making judgments about the impact of the intervention, the theory of change together with data on implementation process and context were used to systematically consider a) whether there were any features of the context that might be responsible for the changes or lack thereof; and b) whether the data on implementation strength were consistent with the observed changes – i.e. substantial implementation in those settings where changes occurred, and weak implementation where no changes were observed. Together, the analyses of impact, process and context allowed the research team to identify those features of the setting that were likely to be associated with successful implementation of this type of intervention. For further information about the evaluation see (Tougher S, the ACTwatch Group et al. 2012).

Realist evaluation to unpack the effects of context

Realist evaluation (Pawson R and Tilley NJ 1997) argues that in order for evaluations to be useful for decision-makers, they must address the questions “what works, how, in what conditions, and for whom”. It uses a number of conceptual approaches, including mid-range theory (a higher level of abstraction than an operational theory commonly specified in a theory of change); and configures changes in outcome as the consequence of both “mechanism” and context. It uses mixed methods, with quantitative methods usually used to measure effectiveness, and then explanations for changes formulated as a series of context-mechanism-outcome (CMO) conjectures which are then examined for patterns. Realist evaluation does not produce universally valid findings, but rather incremental evidence about what conditions are needed or an intervention to work, and how it works, leading to an “accumulation of insights” to guide decision-makers (Marchal et al. 2012). A recent systematic review identified 18 empirical applications of realist evaluation in health systems research, across both high-income settings and LMICs (Marchal et al. 2012). The authors identified a number of advantages of realist evaluation methods including the ability to focus enquiry into complex phenomena by developing mid-range theories at different levels of the system (individual, group, organisation, society). Two papers in the recent Health Research Policy and Systems issue on systems thinking in HSR use realist evaluation methods to examine the effectiveness of health system leadership and management strengthening (Kwamie A, van Dijk H et al. 2014, Prashanth N, Marchal et al. 2014).

6. Challenges and lessons: What can education systems research learn from health systems research?

Health systems research has developed as a coherent and identifiable field of research, engaging both researchers and policymakers in generating evidence about how health systems can be improved. Yet it is still relatively underfunded, as highlighted in the introduction, with research

fundors much more likely to invest in developing new interventions than in new knowledge about how to strengthen the systems underpinning them.

One challenge has been the difficulty of demonstrating the impact of health systems strengthening efforts on health outcomes. This is partly due to the complex nature of health system reforms discussed above, together with fact that health outcomes are influenced by multiple factors, only some of which lie within the health care system. This places health system researchers in the difficult position of choosing whether to try to measure final health outcomes (often requiring very large sample sizes), with the risk that downstream factors may also affect these; or measure intermediate outcomes such as service use, relying on evidence from other studies to infer an intervention's impact on final outcomes.

More generally, there has been continued debate about rigour in health systems research (see for example, (Mills, Gilson L et al. 2008, Sheikh K, Gilson L et al. 2011). Part of the challenge is presented by the inherent multidisciplinary of the field, and there is a need for far better understanding of how the constructs of internal validity (bias) and external validity (representativeness and generalizability) are applied in quantitative and qualitative research and, of course, the limitations of each mode. But there is another tension around how questions are formulated: one cannot help but be struck by the fact that the HSR studies that are published in the highest impact disciplinary journals (especially economics journals) tend to focus on very narrowly defined questions and on impact evaluations of policy interventions with relatively short causal chains, for example, the impact of community accountability mechanisms on mortality (Bjorkman M and Svensson J 2009); or the impact of removing user fees on health status (Powell-Jackson, Hanson et al. 2014). The research that is needed to understand how these interventions can be sustained and scaled up through health systems would probably be very difficult to publish in the highest impact journals; yet is vital that such research is undertaken, applying the appropriate standards of rigour, and that it can be published in outlets which serve the career concerns of the researchers willing to engage with such questions.

Finally, generalizing results across settings has been another challenge, as the effects of health system reforms are likely to be highly context-specific. This is partly a reflection of funding availability which restricts much health system research to single case studies. However, there is a need for more comparative research on how reforms work across settings and for the development of frameworks which will help to synthesise the relevant features of policy processes and contexts. This is messy research; it may be difficult to attract certain types of scholar to study these systems-level questions.

In this review I have tried to take a broad scoping perspective to identify lessons from health systems research that could inform the development of the emerging field of education systems research. Further debate and engagement with education systems researchers will certainly refine and extend this list, however as a starting point, I suggest the following as the main conclusions from the review:

1. Education systems research should develop as a “question driven field”, admitting a variety of research methods and a range of research questions that encompass both exploratory/descriptive questions as well as evaluative questions. The questions will have to move quickly from taking a diagnostic perspective to one which is shaping policy options; but the endpoint of systems research will not be simply “what works” in education reform, but how can effective strategies be scaled up and sustained, politically, institutionally and financially.

2. No single education system model will drive enquiry. As the field of health systems research has demonstrated, there is a need for different models to address different questions and from a variety of disciplinary perspectives. Models of sub-systems will be as important in informing research as “whole system” models. What distinguishes system research is its recognition of the linkages and interactions among system components, the variety of forms of signal that induce change (incentives, information, accountability) and the responses (intended and otherwise) that these changes provoke.
3. The field of education systems research will have to draw on a broad range of social science disciplines (economics, politics, anthropology, psychology) to understand the nature of the “software” of systems (people, their motivation and their values and norms). It will also need to draw on different disciplines’ understandings of the development and reform of public institutions – e.g. Historical institutionalism to understand the comparative history of health / education system development; public administration to understand government capacity to manage. With multiple disciplines contributing to the field, investments will be needed in communicating across disciplines to understand their individual contributions to method and work with the relevant standards of rigour.
4. Evaluations of education systems interventions should recognise their complexity, and include methods which permit an understanding of how and why interventions work, not just whether they work. Systems research should therefore be strongly theory-of-change led.
5. Education systems researchers should reflect on what structures and organisations might be needed to “build the field”, supporting the development and legitimization of education systems research as a field of academic enquiry. The experience of health systems research, which has developed over time to be supported by journals, conferences, training programmes, is informative; yet there continue to be challenges of securing adequate funding for this type of research, debates about its rigour, and continued need to demonstrate impact.

References

- Alliance for Health Policy and Systems Research (2007). "What is health policy and systems research and why does it matter?" Briefing note 1.
- Alliance for Health Policy and Systems Research (2012). Changing mindsets: Strategy on health policy and systems research. Geneva, Alliance for Health Policy and Systems Research.
- Balabanova D, Mills A, Conteh L, Akkazieva B, Banteyerga H, Dash U, Harmer A, Ibraimova A, I. Z, Kidanu A, Koehlmoos TP, Limwattananon A, Muraleedharan VR, Murzalieva G, Palafox B, Panichkriangkrai A, Patcharanarumol W, Penn-Kekana L, Powell-Jackson T, Tangcharoensathien V and McKee M (2013). "Good Health at Low Cost 25 years on: lessons for the future of health systems strengthening." The Lancet **381**: 2118-2133.
- Bishai D, Paina L, Qingfeng L, Peters DH and Hyder AA (2014). "Advancing the application of systems thinking in health: why cure crowds out prevention." Health Research Policy and Systems **12**(28).
- Bjorkman M and Svensson J (2009). "Power to the people: evidence from a randomized experiment on community-based monitoring in Uganda." Quarterly Journal of Economics **142**(2): 735-769.
- Blaauw D, Erasmus E, Pagaiya N, Tangcharoensathien V, Mullei K, Mudhune S, Goodman C, English M and Lagarde M (2010). "Policy interventions that attract nurses to rural areas: a multicountry discrete choice experiment." Bulletin of the WHO **88**: 350-356.
- Blanchet K, Palmer J, Palanchowke R, Boggs D, Jama A and Girois S (2014). "Advancing the application of systems thinking in health: analysing the contextual and social network factors influencing the use of sustainability indicators in a health system – a comparative study in Nepal and Somaliland." Health Research Policy and Systems **12**(46).
- Castro-Leal, F., J. Dayton, L. Demery and K. Mehra (2002). "Public spending on health care in Africa: do the poor benefit?" Bull World Health Organ **78**(1): 66-74.
- Colclough C, Ed. (1997). Marketizing Education and Health in Developing Countries: Miracle Or Mirage? Oxford, Oxford University Press.
- de Moura Castro C and P. Musgrove (2002). On the Non-existence of "The Social Sector" or Why Education and Health Are More Different than Alike, Draft.
- Elloker S, Olckers P, Gilson L and Lehmann U (2013). Crises, Routines and Innovations – the complexities and possibilities of sub-district management. South African Health Review 2012/3. .
- Padarath A and English E. Durban, Health Systems Trust.
- Frenk J (1994). "Dimensions of health system reform." Health Policy **27**: 19-34.
- Gilson L (1988). Government Health Care Charges: Is Equity Being Abandoned? London, Evaluation and Policy Centre, London School of Hygiene and Tropical Medicine.
- Gilson L (2012). Health systems and institutions. Health systems in low- and middle-income countries; An economic and policy perspective. Smith RD and Hanson K.
- Gilson L, Elloker S, Olckers P and Lehmann U (2014). "Advancing the application of systems thinking in health: South African examples of a leadership of sensemaking for primary health care." Health Research Policy and Systems **12**(30).
- Gilson, L., C. Bowa, V. Briljal, J. Doherty, I. Antezana, M. Daura, M. Mabandhla, F. Masiye, S. Mulenga, C. Mwikisa, S. Mbatsha, D. McIntyre, S. Thomas, S. Lake, K. Odegaard and S. Bennett (2000). The Dynamics of Policy Change: Lessons from Health Financing Reform in South Africa and Zambia. Bethesda, MD, Partnerships for Health Reform, Major Applied Research 1 Technical Paper No. 3, Abt Associates.
- Gorsky, M. (2013). The Idea of a 'Health System': International Health Organisations and the Genesis of Comparative Health Systems Research, 1891-1969. Conference Proceedings: International Health Organizations (IHOs) and the history of Health and Medicine, c.1870-2012, Shanghai.
- Kaler A and Watkins S (2001). "Disobedient distributors: street-level bureaucrats and would-be patrons in community-based family planning programs in rural Kenya. 32: 254–69." Studies in Family Planning **32**: 254-269.

Kwamie A, van Dijk H and Agyepong I (2014). "Advancing the application of systems thinking in health: realist evaluation of the Leadership Development Programme for district manager decision-making in Ghana." Health Research Policy and Systems **12**(29).

Leroy, J. L., H. JP, G. Pelto and S. M. Bertozzi (2007). "Current priorities in health research funding and lack of impact on the number of child deaths per year." American Journal of Public Health **97**(2): 219-223.

Marchal B, Van Belle S, De Brouwere V and S. Witter (2013). "Studying complex interventions: reflections from the FEMHealth project on evaluating fee exemption policies in West Africa and Morocco." BMC Health Serv Res **13**: 469.

McKewan PJ (2014). "Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments." Review of Educational Research.

Mills, A. (2011). "Health policy and systems research: Defining the terrain; identifying the methods." Health Policy and Planning.

Mills, A., Gilson L, K. Hanson, Palmer N and M. Lagarde (2008). " Commentary: What do we mean by rigorous health system research? ." The Lancet **372**(9649): 1527-1529.

Molyneux CS, Atela M, Angwenyi V and Goodman C (2012). " Community accountability at peripheral health facilities: a review of the empirical literature and development of a conceptual framework." Health Policy and Planning **27**(7): 541-554.

Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, Moore L, O'Cathain A, Tinati T, Wright D and Baird J (2015). "Process evaluation of complex interventions: Medical Research Council guidance." British Medical Journal **350**.

Moreno-Serra, R. and P. C. Smith (2015). "Broader health coverage is good for the nation's health: Evidence from country-level panel data." Journal of the Royal Statistical Society A.

Musgrove, P. (2003). "Judging Health Systems: Reflections on WHO's Methods." The Lancet **361**: 1817-1820.

OECD (2004). Towards high performing health systems: Summary report. Paris, OECD.

OECD (Various). PISA In-Depth reports, <http://www.oecd.org/pisa/pisaproducts/pisain-depthreports.htm>.

Pawson R and Tilley NJ (1997). Realistic evaluation. London, Sage.

Powell-Jackson, T., K. Hanson, C. J. Whitty and E. K. Ansah (2014). "Who benefits from removing user fees for health care? Evidence from a randomized experiment in Ghana." Journal of Development Economics **107**: 305-319.

Prashanth N, B. Marchal, Devadasan N, Kegels G and Criel B (2014). "Advancing the application of systems thinking in health: a realist evaluation of a capacity building programme for district managers in Tumkur, India." Health Research Policy and Systems **12**(42).

Pritchett, L. H. (2013). The rebirth of education: Schooling ain't learning. Washington DC, Centre for Global Development.

Pritchett, L. H. and Viarengo M (2009). The illusion of equality: the educational consequences of blinding the weak states. Washington DC, Center for Global Development.

Roberts MJ, W. Hsiao, P. Berman and M. Reich (2008). Getting health reform right: A guide to improving performance and equity. New York, Oxford University Press.

Sheikh K, A. George and Gilson L (2014). "People-centred science: strengthening the practice of health policy and systems research." Health Research Policy and Systems **12**(19).

Sheikh K, Gilson L, Agyepong IA, Hanson K, Ssengooba F and S. Bennett (2011). "Building the Field of Health Policy and Systems Research: Framing the Questions." PLoS Med **8**(8): e1001073.

Smith, R. and K. Hanson (2011). What is a 'health system'. Health systems in low- and middle-income countries: An economic and policy perspective. S. RD and K. Hanson. Oxford, Oxford University Press.

Tougher S, the ACTwatch Group, Ye Y, Amuasi JH, Kourgueni IA, Thomson R, Goodman C, Mann AG, Ren R, Willey BA, Adegoke CA, Amin A, Ansong D, Bruxvoort K, Diallo DA, Diap G, Festo C, Johaness B, Juma E, Kalolella A, Malam O, Mberu B, Ndiaye S, Nguah AB, Seydou M, Taylor M, Torres Rueda S,

Wamukoya M, Arnold F and H. K (2012). "Effect of the Affordable Medicines Facility—malaria (AMFm) on the availability, price, and market share of quality-assured artemisinin-based combination therapies in seven countries: a before-and-after analysis of outlet survey data." The Lancet.

Varghese J, Kutty V, Paina L and Adam T (2014). "Advancing the application of systems thinking in health: understanding the growing complexity governing immunization services in Kerala, India." Health Research Policy and Systems **12**(47).

Varvasovszky Z and Brugha R (2000). "How to do (or not to do) a stakeholder analysis." Health policy and Planning **15**(3): 338-345.

Waddington C and N. K. A. Enyimayew (1990). "A price to pay, part 2: The impact of user charges in the Volta region of Ghana." International Journal of Health Planning and Management **5**(4): 287-312.

Williams, A. (2001). "Science or marketing at WHO? A commentary on 'World Health 2000'." Health Economics **10**(2): 93-100.

Woessmann L (2007). "International evidence on school competition, autonomy, and accountability: A review." Peabody Journal of Education **82**(3): 473-497.

World Health Organization (2007). Everybody's business: Strengthening health systems to improve health outcomes, a framework for action (Draft 7 April 2007). Geneva, World Health Organization.

World Health Organization (2010). Increasing access to health workers in remote and rural areas through improved retention: Global policy recommendations. Geneva, World Health Organization.

World Health Organization (2000). The World Health Report 2000. Health systems: improving performance. Geneva, World Health Organization.

Xu, K., D. B. Evans, G. Carrin, A. M. Aguilar-Rivera, P. Musgrove and T. Evans (2007). "Protecting Households From Catastrophic Health Spending." Health Affairs **26**(4): 972-983.