Methodological reflections on using realist evaluation in a study of fee exemption policies in West Africa and Morocco

Bruno Marchal, Karen Van der Veken, Amina Essolbi, Jean-Paul Dossou, Fabienne Richard and Sara Van Belle

December 2013

SEVENTH FRAMEWORK PROGRAMME

Acknowledgments

The research relating to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-13) under grant agreement no 261449, under the FEMHealth project. FEMHealth is a research consortium including the following partners:

- University of Aberdeen (Immpact), UK (coordinator)
- Agence de Formation, de Recherche et d'Expertise en Santé pour l'Afrique (AFRICSanté), Burkina Faso
- Centre de Recherche en Reproduction Humaine et en Démographie (CERRHUD), Benin
- Ecole National de Santé Publique, Morocco
- Institute of Tropical Medicine, Belgium
- Institut de Recherche en Sciences de la Santé (IRSS), Burkina Faso
- London School of Tropical Medicine, UK
- MARIKANI, Mali

For further details, please see: www.abdn.ac.uk/femhealth

We would like to thank Vincent De Brouwere (FEMHealth Work Package Leader), Sophie Witter (FEMHealth coordinator) and Jean Macq (FEMHealth Scientific Committee) for their critical review of the first draft, as well as Alec Cumming (project manager) for his editing.

Table of Contents

Executive summary

User fee exemption for delivery and emergency obstetric care is a policy that has recently been introduced by a large number of countries, particularly in Africa, with the aim of enhancing access to care and improving maternal and neonatal outcomes. In 2011, the FEMHealth project was established with EC funding to conduct multi-disciplinary evaluations of fee exemption policies in Benin, Burkina Faso, Mali and Morocco. One of its objectives was to develop new methodological approaches for the evaluation of complex interventions in low-income countries.

We considered fee exemption policies, and particularly their implementation, as complex in nature. We adopted the realist evaluation approach, since we aimed at researching the policy adoption and implementation, the perceptions of the health service managers and providers, the mechanisms underlying adoption of the policy and the influence of the context.

In this paper, we present methodological reflections on using realist evaluation in a study of the adoption of fee exemption policies carried out by a multicentre consortium. We focus specifically on the issue of elicitation of the initial hypothesis (the middle range theory) underlying the policy implementation, the issue of mechanisms and the challenges of implementing a realist evaluation in a multicentre consortium.

In eliciting the initial theory on policy adoption, we encountered a number of challenges. These include the choice of the theories that inform the initial middle range theory and the issue of drawing lessons from previous research on policy adoption carried out in quite different contexts. We also noted that the users of services and the community are conspicuously absent in the above described policy models and theories, although their role in policymaking and in keeping actors accountable should be beyond dispute. Fourth, we had some problems identifying mechanisms that may play a role in the transitions between the different levels. This relates to the lack of in-depth studies on the mechanisms underlying policy adoption and policy implementation in the study countries, but perhaps also to the difficulty of defining what constitutes a ‘mechanism’.

In realist literature, there is as yet no consensus on interpretation of the concept of ‘mechanism,’ even though it is a central element in the realist mode of explanation. We present insights from our study of the concept of ‘mechanism’ in political science, sociology and public administration, and show how the categories of situational, action formation and transformational mechanisms help in the analysis of FEMHealth data.

Finally, we discuss challenges we met in the design and implementation of this study within a research consortium in which 4 institutions in the south and 3 institutions in the north participated. There is little documented guidance on how to carry out a realist evaluation, which demands a specific approach to introduce the realist principles to teams of researchers not accustomed to the approach. We took a hands-on approach, building pragmatically on the existing competencies and expertise of the study country teams, and this proved to be an effective, albeit time-consuming approach.

Introduction

User fee exemption for delivery and emergency obstetric care (EmOC) is a policy that has recently been introduced by a large number of countries, particularly in Africa, with the aim of enhancing access to care and improving maternal and neonatal outcomes (De Brouwere et al., 2010, Ridde, 2011). The free caesarean section (CS) policy in Mali was introduced in 2005. It is applied nationally to all caesarean sections in the public sector, and in theory covers all facility-based costs (but not transport). In a three-way partition of costs, families are intended to fund the journey into the health centres, while communities fund the onward referral transport costs, and the state covers the costs of service provision, including accommodation, surgery, laboratory tests, and treatment of complications such as pre-eclampsia and ruptured uterus. Burkina Faso introduced a policy in 2006 that subsidised health facilities for 80% of the cost incurred for normal deliveries and caesarean sections. This policy followed several other programmes introduced by the Ministry of Health to improve care for pregnant women. In Morocco, the fee exemption policy initiated at the end of 2008 abolished user fees in hospitals for normal and complicated deliveries (including in theory in university hospitals if the woman is referred from a second level public facility), resuscitation, transport to the appropriate level, and care for mother and newborn as long as they stay in the facility. It was part of a broad action plan for the health sector, which also included a programme to improve supply of drugs, a health workforce plan and interventions aimed at improving transfers of patients between health facilities. In Benin, the policy introduced in 2009 was more selective, covering caesarean sections only and reimbursing health facilities with a flat fee for each intervention carried out.

In 2011, the FEMHealth project was established, with EC funding, to conduct multi-disciplinary evaluations of fee exemption policies in these four countries (see www.abdn.ac.uk/femhealth). A scan of the literature shows that the number of studies or evaluations of such policies is rising (Ridde and Morestin, 2011, McPake et al., 2011, Richard et al., 2010, Richard et al., 2013). These focus on policy effectiveness in terms of utilisation (De Allegri et al., 2011, Dhillon et al., 2012, Lagarde et al., 2012, Witter et al., 2010), equity (El-Khoury et al., 2012) or cost-effectiveness (Witter et al., 2010). Other studies focus on implementation issues (Ben Ameur et al., 2012, Ridde and Diarra, 2009, Idd et al., 2013, Nimpagaritse and Bertone, 2011, Witter et al., 2012, Witter et al., 2007a) or barriers and facilitators (Ridde et al., 2010). Some studies focus on financing (Witter and Adjei, 2007), or assess the effects of such policies on the health workforce (Carasso et al., 2012, Witter et al., 2007b) or health facilities (Witter et al., 2011). Others still analyse the policy formulation process (Meessen et al., 2011, Witter et al., 2013a). However, few of these studies are explicitly based on a hypothesis, a framework or a theory that would provide a basis for analysis or comparison of such policies. Exceptions include (Hercot et al. (2011), Gilson and McIntyre (2005), Walker and Gilson (2004), Ridde and Diarra (2009), Witter (2009) and Robert et al. (2012).

FEMHealth started from the perspective that these are complex policies, which therefore require tailored evaluation methodologies (Marchal et al., 2013). One of the objectives was to develop new methodological approaches for the evaluation of complex interventions in low-income countries. The importance of complexity for health care policy-making and interventions, as well as research and evaluation, is now acknowledged (Gilson, 2012). However, in the policy and health systems

research (HPSR) literature, conceptual confusion is reflected by the interchangeable use of terms such as ‘complicated’ and ‘complex’ and divergent definitions of what makes a problem, an intervention or a specific setting complex (Marchal et al., 2013). Similar problems affect discussions on what constitutes good designs for evaluation or research of complex interventions (Anderson, 2008, Kernick, 2006).

In FEMHealth, we adopted the realist evaluation approach for a sub-study on policy implementation, a typically complex problem. In this paper, we present some methodological reflections on using realist evaluation in a study of the adoption of fee exemption policies carried out by a multicentre consortium. We focus specifically on the issue of elicitation of the initial hypothesis – or the initial middle range theory – underlying the policy implementation, the issue of mechanisms and the challenges of implementing a realist evaluation in a multicentre study.

Background

A study focusing on policy adoption and implementation

In the domain of maternal and reproductive health, implementation of health policies is now firmly recognised as an important issue: many papers and studies present best practices and evidence of effectiveness, noting that going to scale is hampered by absorption capacity constraints that are related to the health workforce (Koblinsky et al., 2006), infrastructure and accessibility of services in general. While in the past, the need for differentiation of strategies, for instance to reduce maternal mortality (Koblinsky et al., 1999), has been demonstrated, there is little (documented) attention for how the gap between global and national policies on one hand, and the actual implementation at service delivery level on the other hand can be bridged. Even just trying to assess the degree of implementation is a difficult enterprise (Witter et al., 2013b).

Even fewer studies started from an explicit theory or theoretical framework on policy adoption or implementation. Indeed, just a few papers attempt to explain the success or failure of the policy implementation on the basis of theories or models from political science or other social sciences disciplines. For instance, only Ridde and Diarra (2009) and Walker and Gilson (2004) used the concept of street-level bureaucracy (Lipsky, 1980) to explain the role of the providers in the policy implementation gap, despite the central role of this theory in the field of policy implementation (see below). In general, few studies focus on the micro-level of implementation and take the perspective of the implementers (exceptions include Agyepong and Nagai, 2011). This is surprising, as policy implementation has been the subject of study in political science since the 1970s. Numerous authors developed theories and models on the basis of research and evaluation, mainly in high-income countries.

In our study of policy implementation, we examined the implementation gap from the perspective of the local health system managers and the providers. This was explored on the basis of concepts from political science, and more specifically the body of work on policy implementation.

Objectives and research questions of the study

The general objective was to conduct a series of realist case studies in selected districts and compare cases to analyse the policy uptake or adoption.

This objective was translated into the following research questions:

- How is the policy being implemented at the district level and to what degree?
- What are the perceptions of the health service managers and providers in terms of challenges related to this policy change?
- What are the mechanisms that explain uptake and implementation of the policy?
- Which context elements facilitate the adoption of the policy?

Methodology

Aiming to develop a study that would be based on existing theory and that would contribute to theory development (or at least testing), we adopted the realist evaluation approach (Pawson and Tilley, 1997).

Theory-driven inquiry

Realist evaluation is one of the 3 main schools of theory-driven inquiry, developed by Pawson and Tilley (1997). These authors argued that in order to be useful for decision-makers, evaluations need to indicate what works, for whom, in what circumstances, in what respects, over which duration and why, rather than respond to "Does it work?". Realistic evaluation (RE) differs from theories of change (TOC) (Connell and Kubitsch, 1998, Fulbright-Anderson et al., 1998) and theory-driven evaluation (Chen and Rossi, 1983, Chen, 1990) mainly in its philosophical foundations in realism and its view on causality (Connelly, 2007, Mingers, 2000). The TOC approach was developed by the Roundtable on Community Change of the Aspen Institute (Connell et al., 1995, Connell and Kubitsch, 1998). Pragmatic in approach and oriented towards stimulating practical change, TOC was initially used to evaluate community-based programmes that typically involved many actors and intervened at several levels.

Theory-driven evaluation focuses not only on the implementation of the intervention and its effectiveness, but also on the underlying causal mechanisms and the contextual factors that underlie change (Chen, 1990). Theory is defined by Chen and Rossi (1983) as the "*prosaic theories that are concerned with how human organizations work and how social problems are generated*".

Realist evaluation

Its explicit philosophical foundations and its methodology set realist evaluation apart from other theory-driven approaches. Pawson and Tilley refer to scientific realism as their philosophical source of inspiration. This school of realism shares a number of elements with critical realism. It likewise accepts that there is a reality independent of the researcher (natural realism), but that knowing this reality through science is unavoidably relative to the researcher (relativist epistemology). Furthermore, realists consider causality to be generative in nature. In other words, they believe that situations or outcomes observed through research need (and can) be explained by the underlying mechanisms of change. These are the processes, triggered by the intervention, that bring actors to do something, or not, Realists won't be content with demonstrating a constant conjunction between an outcome and an intervention, which in positivist thinking suffices to demonstrate causality (Pawson and Tilley, 1997).

Realists thus emphasise the role of actors in change (agency), but also consider that structure matters, and that there is a permanent interplay between individuals and the institutions, culture and other structural elements of their context. The social reality in which a programme or intervention is embedded is, indeed, made up by the social norms, rules and values, and by the relationships between the people involved (Pawson and Tilley, 1997). Others describe these context factors as the historical, institutional, structural and other factors that constrain the choices of the actors. It includes the backgrounds, experiences, loyalties, cultural expectations, or in general the cultural, social, and economic circumstances of these actors (Dahler-Larsen, 2001).

If all human action is embedded within such a wider range of social processes and interactions, then causal mechanisms reside in social relations and context as much as in individuals. Finding and demonstrating mechanisms is not easy, not least because they are not visible. Furthermore, their importance and effect upon the individual will vary and be influenced by the context. For instance, management strategies based on performance-based financing call upon the motivational driver of economic gain. Another example is professional motivation of nurses. The strength of the effect of an economic incentive will vary in function of context factors such as pre-existing salary levels, the cost of living, etc. However, if such drivers of behaviour are not directly visible and likely to be variable, they can nevertheless be studied and understood.

This view on causality leads realists to consider that policies or programmes do not change a situation or solve a problem, but that they merely provide the actors with resources and opportunities. A successful policy or programme triggers mechanisms among the actors, who act and change (or not) a situation, whereby context conditions matter. An effective intervention will have a mechanism that matches its context (Pawson and Tilley, 1997, p. 71). This leads realists to say that not interventions or policies need to be replicated, but rather that a way must be found to trigger the mechanism that will lead to the desired outcome. Hence the importance of understanding the configuration of intervention-outcome-context-mechanism. "*A mechanism is not a variable but an account of the behavior and interrelationships of the processes which are responsible for the change. A mechanism is thus a theory*" (Pawson and Tilley, 1997). This makes realist evaluations almost by default a multi-disciplinary endeavour, since mechanisms have been described in the social sciences, political science, history, economics, psychology, organisational theory, etc.

General design

We adopted the comparative case study design, which is well suited for research of the policy-implementation gap. This design allows exploring a "*phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident*" (Yin, 2003 p. 13) and is indicated for studies of small group behaviour and organisational processes (Yin, 2009).

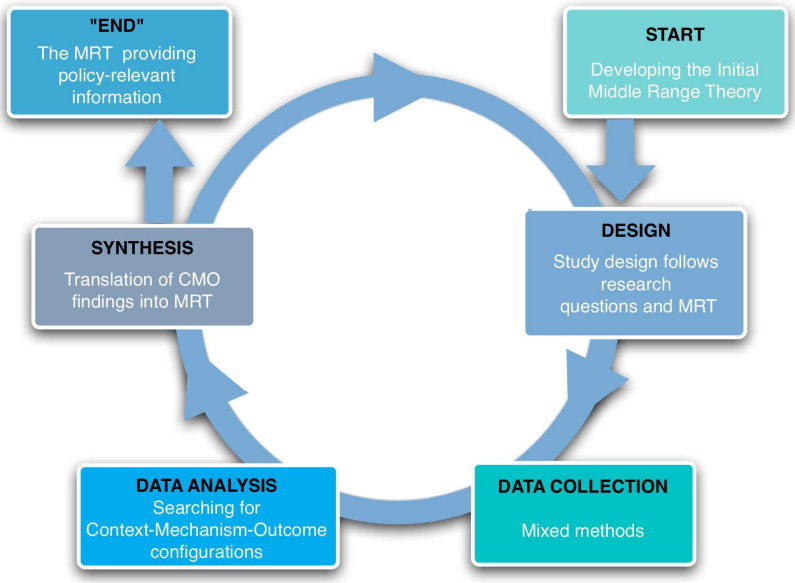
In Morocco, Benin and Burkina Faso, two districts were selected among the FEMHealth sites in April 2013 on the basis of a previous study on the effects of the policy at the local health system level (see www.abdn.ac.uk/femhealth/about /programme-outputs). The civil war that erupted in Mali in 2011-12 precluded the planned field research and the Mali sites therefore had to be dropped in this study. In line with realist evaluation principles, the case selection was purposive. To investigate the gap between policy and implementation, we selected one district with strong implementation and one with weak implementation of the policy, and more specifically with apparent differences in policy adoption by health service managers

and providers. Data were collected through in-depth interviews and document review. Key respondents included health service managers (district health management team, hospital management team and mid-level cadres), and health care providers at hospital level. In line with RE principles, the data analysis was based on the core elements of the middle range theory, leaving space for emerging insights to inform the analysis. The final analysis is yet to be concluded at the time of writing of this paper, but we present some preliminary results in the last but one section of this paper.

Realist evaluation in practice

Realist evaluation starts and ends with theory (Figure 1), which needs to be understood as theories of the middle range: "theories that lie between the minor but necessary working hypotheses (…) and the all-inclusive *systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization and social change*" (Merton, 1968). The initial middle range theory (MRT) states how the intervention is expected to lead to its effect and in which conditions. It can be formulated on the basis of existing theory, past experience, previous evaluations or research studies and insights and perceptions of the involved actors.

Figure 1 – The realist cycle, adapted from (Marchal et al., 2012).



In a next step, the design of the study is chosen, and data collection tools are developed and tested. Realist evaluation is method-neutral: the study design and data collection methods should allow 'testing' of the key elements and assumptions. Often both quantitative and qualitative data are collected, according to the focus of the MRT. The data collection phase is followed by data analysis, whereby realist evaluation uses the context-mechanism-outcome (CMO) configuration as the main heuristic. The resulting CMO configurations are then assessed to see whether they hold as plausible patterns (in realist evaluation also called 'demi-regularities') that explain how the intervention brought about the observed results. These CMOs are

finally compared with the initial middle range theory, which is then refined. This kicks off the next study in a cycle that refines the theory through accumulation of better insights (Pawson and Tilley, 1997, Marchal et al., 2012).

How to elicit the middle range theory

As mentioned above, a middle range theory is the starting point of a realist evaluation. We present briefly different approaches to how a MRT can be elicited as proposed by Lipsey and Pollard (1989), adding insights from other authors and from studies in theory-driven evaluation.

A first method is to examine existing knowledge and theories. In some cases, the problem situation or the intervention or the policy under study has been researched extensively, allowing formulation of an MRT on the basis of past experience and existing theory. In other cases, appropriate concepts can be searched for in disciplines such as psychology, sociology and other social sciences. Approaches for meta-review, such as meta-narrative review (Greenhalgh et al. 2005), realist synthesis (Pawson et al. 2005) and meta-triangulation (Lewis and Grimes, 1999) are well suited (see also Greenhalgh et al. 2011). Leeuw (2003) demonstrates how argumentational analysis can be used to draw useful propositions from data sources such as stakeholder interviews, project documents and reports and published documentation. Since many policy and project documents lack an explicit and well-articulated argumentation of why and how the expected goals can be achieved, the researcher needs to 'mine' the available information. Argumentational analysis is a "*model for analysing chains of arguments and it helps to reconstruct and 'fill in' argumentations*" (Leeuw, 2003). Applied to realist evaluation, this technique can be used to search the data for statements that point to expected mechanisms and context conditions. These resulting statements are listed and the mechanisms, outcomes and context conditions they refer to are linked into configurations. These are in turn analysed and compared with existing research findings and theories. When little theoretical knowledge is available, a relevant MRT could be drafted by reversing the theory that explains the problem tree (problem analysis) to describe potential theory about the solution.

A second approach consists of exploratory on-site research in an effort to build an initial MRT on the basis of empirical study of the problem or intervention of interest (Lipsey and Pollard, 1989). Principled discovery, proposed by Julnes and Mark (1998), is a similar technique. Applied to the analysis of data sets of existing cases, patterns are sought that explain differences in outcomes and these are developed into explanatory theories.

Third, if an intervention has already been designed or when its implementation has started, (elements of) the MRT can be extracted from interviews of the designers and implementers of the intervention (Lipsey and Pollard, 1989, Birckmayer and Weiss, 2000). The researchers unearth the models that the actors are implicitly using to describe and understand the intervention - what Pawson and Tilley (1997) call 'folk theories' - through individual interviews or group discussions. Additional information may be derived from programme documents or policy documents. Techniques such as cause mapping and concept mapping (Rosas, 2005) and argumentational analysis (Leeuw, 2003) may help to clarify how the key actors understand the intervention. This should result in the identification of the key elements of the problem or intervention, and, in the latter case, of the expected short- and long-term outcomes as described by the stakeholders. This is similar to what Trochim (1989) calls 'pattern matching', whereby the stakeholders are interviewed and the pattern of their propositions compared with observations or prior theory..

Once the initial MRT is elicited, the researchers may discuss it with the stakeholders and check the literature to assess its plausibility.

Our initial middle range theory on policy implementation

We started the MRT development process by scoping the literature on fee exemption for maternal health services. We aimed to find out which theories or frameworks are used to assess the adoption and implementation of fee exemption policies in this literature. In essence, this review showed that there are few studies that explore how and why policies are not being implemented beyond factors of policy formulation and support. We therefore carried out a review of the policy implementation literature from political science, public administration and policy analysis.

Our aim was to obtain an understanding of the range of theories and frameworks of policy implementation. We specifically intended to tease from this literature relevant categories of context and mechanisms that would explain how particular policies are implemented (or not). Since this literature is quite fragmented and spread over a number of disciplines and domains, with little if any consensus on theories (O'Toole, 2004, DeGroff and Cargo, 2009), an exhaustive and systematic review was not likely to prove useful. We therefore carried out a non-exhaustive review of the published and grey literature, which was inspired by principles of realist synthesis (Greenhalgh et al. 2011) and interpretative in nature. We started the search using the Web of Science/Social Sciences Citation Index search engines with sets of key words combining ‘policy’, ‘program’ and ‘implementation’ or ‘adoption’. We used extensive snowballing to track down the original papers and books from the bibliographic references of the initial list of papers and books. Our search did not claim, therefore, to be exhaustive.

From this literature review, we retained the framework developed by Berman (1978) as it provides an interesting and useful theory to study policy implementation. His framework takes a systemic perspective, usefully linking various elements or levels of what is sometimes called a policy implementation system. The author posits a series of transitions during a macro-implementation phase, which includes the shift from policy decision to programme, from programme to local adoption, and from adoption to implementation. He then describes a micro-implementation phase during which local managers and providers interpret and apply (or not) the programme. The author already pointed to some factors we could consider as mechanisms. For instance, how the programme will be adopted at the operational level will depend on the compliance (and thus enforcement capacity) and the consonance or alignment with local needs and goals. Another pointer to a mechanism during the micro-implementation phase, is the finding that commitment of implementers can be generated through involving them in decision-making and planning.

Another potential set of categories to think about in terms of causal models is provided by O'Toole and Montjoy (1984). If collaboration is at the core of policy implementation, these authors argue that a policy can include 3 types of inducement for cooperation: authority (defined as cooperation derived from a sense of duty), common interest (cooperation on basis of shared values) and exchange (cooperation to obtain other goals). Policies are more likely to be implemented by an organisation if the policy is aligned with the organisational goals and its routines, if additional new resources are provided and if sanctions can be imposed in case of inadequate implementation. When multiple agencies need to collaborate for successful policy implementation, these authors claim that formal authority is required to enforce the implementation, unless the new policy is closely aligned with the goals of these organisations. Pressure is required in most settings to focus attention on a reform objective; support is needed to enable implementation (McLaughlin, 1987).

What is called the policy implementation system (the set of actors, agencies and procedures concerning the policy) should be taken into account: not just the policy or the programme, but the whole process of negotiation and adaptation that occurs at all levels during the implementation needs to be included. Since implementation is best considered as a complex process, " "why" and "how" are as critical as "what" and "how much" " (McLaughlin, 1987). McLaughlin proposes to combine a bottom-up, micro-level analysis with a top-down, macro-level analysis (McLaughlin, 1987).

Similarly, Yanow (1987) proposes 4 lenses to be adopted when analysing policy implementation: human relations (the individual behaviour within organisations), politics (dynamics within groups and relations among them), structure (the organisation as a set of behavioural rules), and system (framing organisations within their larger system). Each of these lenses comes with specific concepts:

- The human relations lens: role definition and expectations, interpersonal skills, motivation, commitment, etc. (psychological factors)
- The political lens: power, influence, interests, coalitions, negotiation, bargaining (political factors)
- The structural lens: design, structure and culture of the organisation (organisational factors) and the relations with other actors (see also DeGroff and Cargo, 2009).
- The systems lens: the organisation as part of a complex adaptive system (systems factors), which may be considered to include networked governance arrangements (see also DeGroff and Cargo, 2009)

For Yanow (1987), these perspectives need to be complemented by information on the policy's history and culture (including the norms and beliefs underlying the policy), and the wider political context. This is what she calls the cultural lens, "*a cultural approach to implementation highlights particular features of implementation: interpretation and shared, compatible, and conflicting meanings, (…) the role of persuasion, and the on-going reinterpretation of implementation activities.*" Such interpretative analysis may reveal that policy implementation is not merely about gaps between stated policies and policy intentions on one hand, and actual implementation and outcomes on the other – which can be due to psychological, political, organisational or systems factors – but also potentially due to inadequate support because of lack of consensus in support of the policy. Yanow's frame thus allows looking at the implementation system in a comprehensive way.

Summary of the key elements of the initial MRT

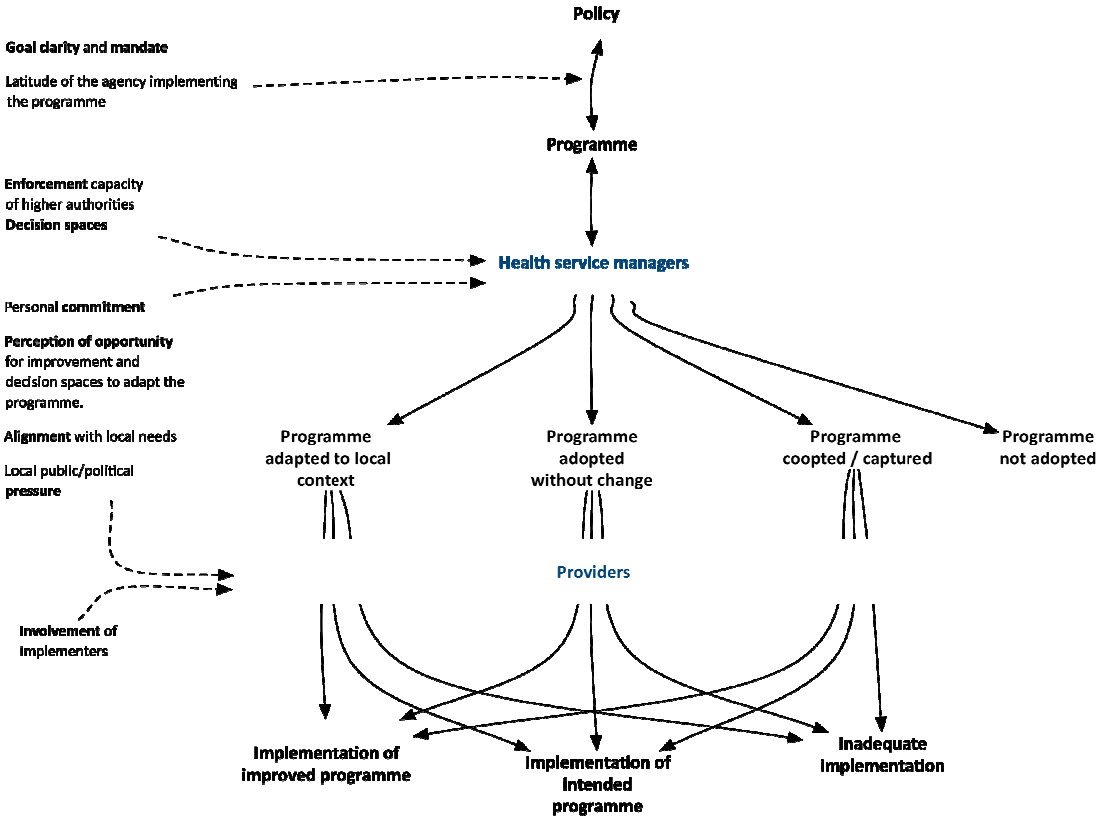
Starting from the approach of Berman (1978), and adding elements of Matland, (1995), Yanow (1987) and O'Toole and Montjoy (1984), we can identify a number of elements that help to formulate our initial MRT.

- The degree of implementation of a health policy depends on the local service managers, the providers and the community, who in their actions are influenced by factors at organisational and local community level as well as at the programme and policy level.
- More specifically, the programme is more likely to be fully adopted by the service providers and service managers if they feel a commitment towards the policy (public motivation/shared values), feel obliged to do so (coercion by their superiors, the authorities or the community) or believe they may gain from it in term of financial or social benefits (social or economic exchange). Facilitating factors include good alignment of the policy with personal goals and with organisational goals and culture, adequate resourcing and organisational support, and an effective monitoring and sanctions system (including local

political pressure or pressure from the public). The latter in turn requires that community representatives are well informed on the policy and have effective 'voice' and 'participation' channels. If such facilitators are absent, capture by elites is more likely.

- The programme is likely to be adapted in a positive sense if service providers and managers believe it can be improved. Besides motivated staff, this requires adequate capacity and decision spaces to adapt the programme to the local situation.
- Adoption by service managers and providers will be enhanced if the translation of the policy into a programme includes the development of clear operational guidelines, indicating the goals, the target groups and the practical implementation modalities (in this case including the reimbursement mechanisms, supply systems, reporting, etc.), and if all necessary inputs (funding, consumables), training (when necessary) and compensation for lost revenue are provided. A dedicated agency with a clear mandate may facilitate implementation. Besides these technical inputs, processes matter. The degree of consensus on the goals, target groups and implementation modalities (low levels of ambiguity and of conflict) influences the passage from policy to implementation.

Figure 2 - The initial MRT



As explained above, this MRT was translated into a set of data collection tools (mainly interview guides for a variety of key respondents) and into an analytical guide.

Methodological reflections

At this stage, the data analysis is being finalised and we cannot therefore present final results. However, we encountered a number of challenges and we present here a number of reflections on using realist evaluation in the study of policy implementation. More specifically, we focus on challenges encountered when eliciting the middle range theory and the issue of what constitutes a mechanism in the field of policy adoption. We then discuss how realist evaluation was applied by a multidisciplinary group of researchers within the FEMHealth project.

Eliciting the initial theory on policy adoption

For the study on policy implementation, and more specifically the adoption of the policy by the service managers and providers, we elicited a middle range theory on the basis of a review of the policy implementation literature. During this process, we encountered several problems.

The first challenge is the choice of the theory. It is clear that despite more than 40 years of research, there is still little consensus in the field of policy implementation studies on the theoretical principles and research methods (O'Toole, 2004). Furthermore, not only the field of policy implementation, but also other approaches need to be taken into account. This is what Pawson (2013a) calls theory adjudication, or deciding which theory holds a better potential of explanation than another. This is, indeed, an issue that is well acknowledged in realist evaluation. Theories are deemed to be useful as long as they help in providing plausible explanations. Realists will be careful not to claim universal applicability of their findings. Indeed, *specification* through repeated 'testing' of the middle range theory is a key process in realist inquiry.

Second, it is striking that most theoretical, methodological and empirical work has been carried out in North America and, to a lesser extent, in the UK or Scandinavia. This specific context matters in the analysis, not in the least because the theories attempt to explain policy making and implementation in particular political systems that are quite different from the governance systems found in the countries where FEMHealth carried out its research. This suggests that we need to ensure that we pay sufficient attention to the broad governance structures and arrangements in the study countries, as well as to the nature of their bureaucracies and other organisations that intervene in the implementation.

Third, the users of services and the community are conspicuously absent in the above described policy models and theories, even though their role in policymaking and in keeping actors accountable is beyond dispute.

Fourth, we had some problems in identifying mechanisms that may play a role in the transitions between the different levels. This relates to the lack of in-depth studies on the mechanisms underlying policy adoption and policy implementation in the study countries, but perhaps also to the difficulty of defining what constitutes a 'mechanism'.

The issue of 'mechanism'

As we mentioned in the Methodology chapter, mechanisms are at the heart of the realist explanation. Mechanisms are not programme activities or modalities, and they are usually not easily visible. Indeed, mechanisms are best understood as underlying entities, processes or structures that lead to outcomes in specific conditions (Astbury

and Leeuw, 2010). In the words of Pawson and Tilley (1997, p. 57), "*programs work (have successful 'outcomes') only in so far as they introduce appropriate ideas and opportunities ('mechanisms') to groups in the appropriate social and cultural conditions ('context')*".

We also saw how mechanisms play out at the level of individuals, groups, organisations and/or society. If a programme is best considered as a set of resources (material, social or cognitive), which actors use or not to change an initial situation (Pawson, 2013b), we can search for mechanisms in psychological, social, cultural, political and economic theories.

The realist literature shows that there is as yet no consensus on the interpretation of 'mechanism' (Marchal et al., 2012). Some authors stick to the definitions provided by Pawson and Tilley (1997). Pommier et al. (2010), for instance, cite from Pawson & Tilley's book: "*A mechanism is not a variable but an account of the behavior and interrelationships of the processes which are responsible for the change. A mechanism is thus a theory*". The definitions presented by Rycroft-Malone et al. (2010), Ogrinc and Batalden (2009) and Clark et al. (2005) seem also to align with Pawson and Tilley's view. Yet, Evans and Killoran (2000) define mechanisms to include interventions or activities, while for Greenhalgh et al. (2009), mechanisms are "*the stakeholders' ideas about how change will be achieved*". The mechanisms they found include descriptions of the actual intervention.

Another main source of diverging views is the difference between mechanism and essential context condition. What Greenhalgh et al. (2009), for instance, call 'conditions for success' could also be considered as 'mechanisms'. They present diagrams that spell out constraining and enabling factors (what others would call context conditions) and 'success' and 'disappointment' (outcomes). Also in the paper by Ogrinc and Batalden (2009) on the effect of teaching on improvement of quality of care, there is some confusion between mechanism and context: is the learners' schedule a mechanism or a context element that needs to be taken into account?

Mechanisms in political science

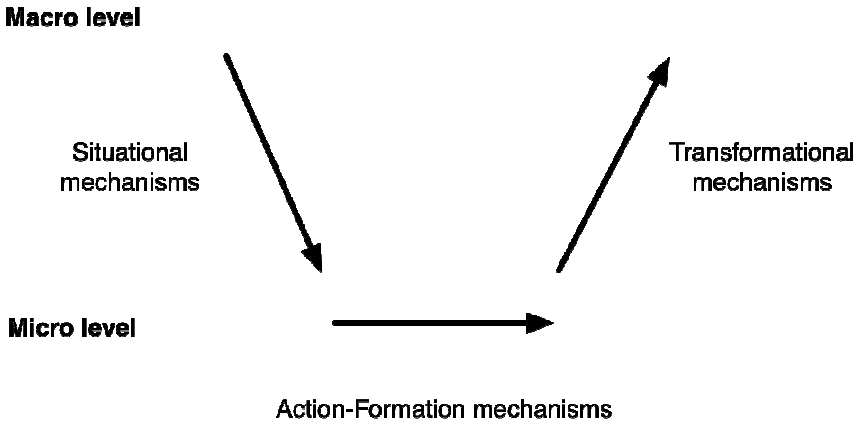
Few policy implementation researchers describe mechanisms in the sense as defined above. O'Toole and Montjoy (1984) come very close when they identify authority, common interest and exchange as the 3 types of inducement for cooperation that policymakers need to induce in order for the policy to be implemented. Also the perspective of Yanow (1987) is useful as her proposed lenses of human relations, politics, structure and system allow looking for mechanisms within each of these dimensions. Hers reflects the agency-structure interaction that realists hold to be underlying generative causation.

In political sciences, there are other streams that are mechanism-based. Hedström and Swedberg (1998), for instance, define mechanisms as "*analytical constructs that provide hypothetical links between observable events*". This is a much broader definition than that usually provided by realists. Similarly, Tilly (2001) distinguishes between three categories of mechanisms to describe political processes: environmental, relational and cognitive mechanisms. Cognitive mechanisms exist at the level of individual and collective perception. Environmental mechanisms are "*externally generated influences on conditions affecting social life*" and relational mechanisms operate at the level of *"connections among people, groups and interpersonal networks*". In this way, Tilly (2001) situates mechanisms also at other levels than the micro level.

Hedström and Swedberg (1998) present yet another typology of social mechanisms, which centres around "*individual action oriented towards others*". The authors distinguish three categories of actions, situational, action-formation, and transformational mechanisms.

- **Situational mechanisms** (macro-to-micro) include the social situations or events that shape the desires and beliefs of individuals. Reference group theory provides examples of such a mechanism. Reference groups provide benchmarks that humans use for comparison and evaluation of group and personal qualities, circumstances, attitudes, values and behaviours. Individuals compare themselves with reference groups of people who occupy the social role to which they aspire, as they act on the conviction that belonging to and remaining part of the reference group requires conformity of beliefs and behaviour. Another example is high-commitment management, where a comprehensive set of synergetic management practices (from adequate remuneration to training and participative decision-making) leads to organisational commitment because these practices trigger reciprocal behaviour.
- **Action formation mechanisms** operate at the level of micro-to-micro, and include the desires, beliefs and opportunities that shape individual action. Here, mechanisms explain how the agency of one individual affects the agency of another individual. Mechanisms can thus be found at the level of intrinsic motivation, for instance, and in cognitive psychology.
- **Transformational mechanisms** (micro-to-macro) describe how individuals shape macro-level outcomes through their actions and relations. This is about how the agency of individuals and small groups shape the structural dimension of society. Examples include the bottom-up influence of practitioners on policymakers engendered by supervision and monitoring, or by performance reviews. Assumed mechanisms could be naming-and shaming of policymakers, awareness raising through confrontation with field realities, pressure exerted on programme managers by operational staff, etc.

Figure 3 – A typology of social mechanisms (Hedström and Swedberg, 1998)



Mechanisms in FEMHealth

If the transition from a policy to a programme and the programme itself is considered to be part of the distal context of a local health system, then mechanisms triggered by these external influences can be classified as *situational mechanisms*. These include the inducements identified by O'Toole and Montjoy (1984), and more specifically 'authority', which works if conditions for enforcement are present, even in situations where the implementers tend not spontaneously to implement the programme. This was translated in our initial MRT as 'coercion'. Also participatory policy-making processes could be considered as mechanisms of this type, inducing ownership and thus a motivation to adopt and implement programmes. Mission valence (in our initial MRT 'the alignment' between the official mission of the organisation - *in casu* the Ministry of Health -, its policy and the personal motivation of the implementers), also matters. The nature and aim of the policy may stimulate (or not) managers and providers to adopt and implement the policy because it is well aligned with their personal values or with their personal or organisational interests. Such alignment may thus trigger social and economic exchange mechanisms respectively.

Action-formation mechanisms operate at the micro-to-micro level. Here, the relationship between service and facility managers and their operational staff matters. The local organisational culture can also be important in steering the response of providers. In addition, the degree of public service motivation of service managers and providers shapes their response to a public policy.

Finally, transformational mechanisms can be triggered if the policy introduction is accompanied by monitoring and evaluation procedures that lead to effective learning form operational experience. If central-level staff, be it in a supportive or monitoring role, visit the operational level, opportunities for exchange and feedback may emerge and feedback may connect local agents with the structural policy level.

Sets of mechanisms identified in our studies

A general pattern that seems to emerge across all study countries is the importance of structural factors and their influence on local agents. Indeed, in most countries, the policy was adopted by the implementers, i.e. both managers and providers, by default; in virtually all sites, the local agents expressed the view that it is not possible to formally go against a central dictum, in this case the policy as imposed by the central administration. Where the bureaucratic system is well developed, perceived coercion seems stronger, although it must be said that the bureaucratic control of the central level over the operational actors is not very strong in all countries.

However, this formal acceptance of the policy in most cases did not lead to pro-active adoption of the policy and its adaptation to local contexts. We found that local context matters: where the local organisational culture is one of laissez-faire and of formal compliance, the policy is implemented on paper and to a minimum degree. In Morocco, for instance, we found cases where the specialists formally adopted the policy, but took a literal reading of the minimum requirements to implement the policy. For example, they used quality criteria included in the policy as a pretext to not organise a permanent call system. The policy was thus formally adopted, but also used to maintain a situation comfortable for the providers but negative for the patients. Such practices were condoned by the administration, who claimed to be powerless vis-à-vis the specialists and who did not wish to disturb the negotiated order with these providers.

In other cases, we found that the local organisational context was favourable to policy implementation, or vice versa that the policy was found to be favourable to the organisational mission – a case of strong mission valence. For instance, in one hospital in Benin, the management team used the generous reimbursements to strengthen other non-targeted services and to better remunerate all staff, not because the policy document said so (which indeed it didn't), but because the management team recognised the opportunity and used it strategically to strengthen the service delivery of the whole hospital.

Examples of how local action shaped the policy formulation at central level were found in Benin, for instance, where the interaction between the agency responsible for programme implementation and monitoring on one hand, and the implementers on the other, allowed for effective support but also for adaptation of the policy to field realities. In this case, at least during the initial phase of policy roll out, the feedback loop was effectively closed, and lessons were learned from initial failure.

To conclude this section on preliminary results, it seems that effective policy adoption and implementation requires triggering mechanisms at all levels, at macro-to-micro, micro-to-micro and micro-to-macro level.

Applying realist evaluation in a multi-centre study

We applied the realist evaluation approach in a study in which 4 institutions in the south (one in each study country) and 3 institutions in the north participated. Each of these partners fielded a research team composed of different types of researchers, including economists, sociologists, anthropologists, midwives, doctors and public health specialists. Applying the realist evaluation approach in such a multi-disciplinary consortium posed some challenges.

First, as already mentioned, realist evaluation is an approach that has been used relatively little in health systems research, and few research centres have a strong experience with this methodology. This raised the issue of familiarising and training the researchers mainly responsible for the data collection and analysis, against a background of little documented guidance on how to do a realist evaluation. We set out to introduce the basic concepts during the consortium's planning meetings and through dialogue with the researchers assigned to this work package. Follow-up was ensured during research meetings and workshops, during which the evolving protocol of the study was presented and discussed. Email contacts and personal discussions during field visits added opportunities for knowledge and skill transfer adapted to the background and experience of each of the teams. This took more time than planned, more so because other aspects of the FEMHealth programme usually had to be discussed and organised at the same time.

Once we managed to get all teams to a shared level of understanding of the principles of realist evaluation, the rest proved much easier. Indeed, realist evaluation is method neutral, and we could therefore appeal on the quantitative and qualitative research skills and competences of the research teams during the phase of data collection design. Developing data collection guides and analytical frameworks was relatively straightforward but time consuming. The same applies to the actual data collection exercises, where we simply tapped into the strong capacity of the research teams to organise and carry out complicated large-scale data collection exercises.

Transmitting the analytical approach required to make sense of data in a realist evaluation to research team members was found to be more challenging. Here, we found that the response of each team member to the realist analytical approach was influenced to a large extent by his/her scientific education or disciplinary background.

Some sociologists, for instance, had no trouble seeing the advantages (and disadvantages) of RE and they quickly grasped the essentials. For some of the public health doctors, RE opened new horizons, providing a new way to approach and analyse health systems problems. Yet, other researchers were less charmed by the approach and preferred sticking to their guns. The same issues popped up during the data analysis, which in this study was largely qualitative in nature. This was not an issue for our qualitative researchers in each of the teams. However, the realist take to analysis of data, and especially the search for the underlying mechanisms that explain the observed patterns was less easy to organise. To this end, a week-long research meeting was held, bringing researchers together to review results and discuss. In all cases, the involvement of a core team of researchers from the ITM in the whole process helped to bridge gaps in the study country teams if and where they occurred.

Conclusions

Adherents of realist evaluation hold that it is a research approach that is well suited for the study of complex issues, such as the implementation and adoption of fee exemption policies. The realist view on causality allows for opening the black box between policy or programme on one hand and outcomes on the other hand, whereby both the context and the response of the actors who are supposed to adopt and implement the policy are taken into account. A key issue in this approach is the mechanism that explains why the policy was adopted and eventually implemented.

We found that the political science literature offers may ways to conceptualise policy implementation and indirectly also a number of mechanisms that can be used to explain why the policy is adopted or not. In this study, we elicited an initial middle range theory on the basis of a review of the policy implementation theory that discerns a series of transitions during the macro-implementation phase, including a shift from policy decision to programme, from programme to local adoption, and from adoption to implementation.

The political science and sociology literature uses definitions and conceptualisation of mechanisms that at first seem to be different from the realist definitions. However, we found that the perspective of Hedström and Swedberg offers a useful frame to examine mechanisms in policy implementation, and our preliminary results show how different mechanisms may need to be triggered at different levels for a policy to ‘work’.

There is little documented guidance on how to carry out a realist evaluation, which demands a specific approach to introduce the realist principles to teams of researchers not accustomed to the approach. We took a hands-on approach, building pragmatically on the existing competencies and expertise of the study country teams, and this proved to be an effective, albeit time-consuming approach.

References

AGYEPONG, I. A. & NAGAI, R. A. 2011. "We charge them; otherwise we cannot run the hospital". Front line workers, clients and health financing policy implementation gaps in Ghana. *Health Policy,* 99**,** 226-33.

ANDERSON, R. 2008. New MRC guidance on evaluating complex interventions. *BMJ,* 337**,** a1937.

ASTBURY, B. & LEEUW, F. 2010. Unpacking black boxes: mechanisms and theory building in evaluation. *American Journal of Evaluation,* 31**,** 363-381.

BEN AMEUR, A., RIDDE, V., BADO, A. R., INGABIRE, M. G. & QUEUILLE, L. 2012. User fee exemptions and excessive household spending for normal delivery in Burkina Faso: the need for careful implementation. *BMC Health Serv Res,* 12**,** 412.

BERMAN, P. 1978. The study of macro and micro implementation of social policy, Santa Monica, CA, RAND Corporation.

BIRCKMAYER, J. D. & WEISS, C. H. 2000. Theory-based evaluation in practice. What do we learn? *Eval Rev,* 24**,** 407-31.

CARASSO, B. S., LAGARDE, M., CHEELO, C., CHANSA, C. & PALMER, N. 2012. Health worker perspectives on user fee removal in Zambia. *Hum Resour Health,* 10**,** 40.

CHEN, H. & ROSSI, P. 1983. Evaluating with sense - The theory driven approach. *Evaluation Review,* 7**,** 283-302.

CHEN, H.-T. 1990. *Theory-driven evaluations,* Newbury Park, California, Sage Publications.

CLARK, A. M., WHELAN, H. K., BARBOUR, R. & MACINTYRE, P. D. 2005. A realist study of the mechanisms of cardiac rehabilitation. *J Adv Nurs,* 52**,** 362-71.

CONNELL, J. P., KUBISCH, A., SCHORR, L. & WEISS, C. 1995. *New approaches to evaluating community initiatives. Concepts, methods, and contexts,* Washington D.C., The Aspen Institute.

CONNELL, J. P. & KUBITSCH, A. C. 1998. Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems. *New approaches to evaluating community initiatives, Vol. 2.* Washington: Aspen Institute.

CONNELLY, J. 2007. Evaluating complex public health interventions: theory, methods and scope of realist enquiry. *Journal of Evaluation in Clinical Practice,* 13.

DAHLER-LARSEN, P. 2001. From programme theory to constructivism. On tragic, magic and competing programmes. *Evaluation,* 7**,** 331-349.

DE ALLEGRI, M., RIDDE, V., LOUIS, V. R., SARKER, M., TIENDREBEOGO, J., YE, M., MULLER, O. & JAHN, A. 2011. Determinants of utilisation of maternal care services after the reduction of user fees: a case study from rural Burkina Faso. *Health Policy,* 99**,** 210-8.

DE BROUWERE, V., RICHARD, F. & WITTER, S. 2010. Access to maternal and perinatal health services: lessons from successful and less successful examples of improving access to safe delivery and care of the newborn. *Tropical Medicine & International Health,* 15**,** 901-9.

DEGROFF, A. & CARGO, M. 2009. Policy implementation: implications for evaluation. *New Directions for Evaluation,* 124**,** 47-60.

DHILLON, R. S., BONDS, M. H., FRADEN, M., NDAHIRO, D. & RUXIN, J. 2012. The impact of reducing financial barriers on utilisation of a primary health care facility in Rwanda. *Glob Public Health,* 7**,** 71-86.

EL-KHOURY, M., HATT, L. & GANDAHO, T. 2012. User fee exemptions and equity in access to caesarean sections: an analysis of patient survey data in Mali. *Int J Equity Health,* 11**,** 49.

EVANS, D. & KILLORAN, A. 2000. Tackling health ineaqualities through partnership working: lessons from a realistic evaluation. *Critical Public Health,* 10**,** 125-40.

FULBRIGHT-ANDERSON, A., KUBISCH, A. & CONNELL, J. 1998. *Theory, measurement, and analysis,* Washington, DC, Aspen Institute.

GILSON, L. 2012. *Health Policy and Systems Research. A methodology reader,* Geneva, Alliance for Health Policy and Systems Research & World Health Organisation.

GILSON, L. & MCINTYRE, D. 2005. Removing user fees for primary care in Africa: the need for careful action. *BMJ,* 331**,** 762-5.

GREENHALGH, T., ROBERT, G., MACFARLANE, F., BATE, P., KYRIAKIDOU, O., PEACOCK, R. 2005. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. Social Science and Medicine, 61, 417-430.

GREENHALGH, T., HUMPHREY, C., HUGHES, J., MACFARLANE, F., BUTLER, C. & PAWSON, R. 2009. How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *Milbank Q,* 87**,** 391-416.

GREENHALGH, T., WONG, G., WESTHORP, G. & PAWSON, R. 2011. Protocol--realist and meta-narrative evidence synthesis: evolving standards (RAMESES). BMC medical research methodology, 11, 115.

HEDSTRÖM, P., SWEDBERG, R., 1998. Social mechanisms: An Introductory essay. *In:* HEDSTRÖM, P., SWEDBERG, R. (ed.) *Social Mechanisms: An Analytical Approach to Social Theory.* Cambridge: Cambridge University Press.

HERCOT, D., MEESSEN, B., RIDDE, V. & GILSON, L. 2011. Removing user fees for health services in low-income countries: a multi-country review framework for assessing the process of policy change. *Health Policy Plan,* 26 Suppl 2**,** ii5-15.

IDD, A., YOHANA, O. & MALUKA, S. O. 2013. Implementation of pro-poor exemption policy in Tanzania: policy versus reality. *Int J Health Plann Manage,* Epub ahead of print.

JULNES, G. & MARK, M. 1998. Evaluation as sensemaking: knowledge construction in a realist world. *New Directions for Evaluation,* 78**,** 33-52.

KERNICK, D. 2006. Wanted--new methodologies for health service research. Is complexity theory the answer? *Fam Pract,* 23**,** 385-90.

KOBLINSKY, M., MATTHEWS, Z., HUSSEIN, J., MAVALANKAR, D., MRIDHA, M., ANWAR, I., ACHADI, E., ADJEI, S., PADMANABHAN, P., MARCHAL, B., DE BROUWERE, V., VAN LERBERGHE, W. & ON BEHALF OF THE LANCET MATERNAL SURVIVAL SERIES STEERING GROUP 2006. Going to scale with professional skilled care. *Lancet,* 368**,** 1377-86.

KOBLINSKY, M. A., CAMPBELL, O. & HEICHELHEIM, J. 1999. Organizing delivery care: what works for safe motherhood? *Bull World Health Organ,* 77**,** 399-406.

LAGARDE, M., BARROY, H. & PALMER, N. 2012. Assessing the effects of removing user fees in Zambia and Niger. *J Health Serv Res Policy,* 17**,** 30-6.

LEEUW, F. 2003. Reconstructing program theories: methods available and problems to be solved. American Journal of Evaluation, 24, 5-20.

LEWIS, M. W. & GRIMES, A. J. 1999. Metatriangulation: Building Theory From Multiple Paradigms. Academy of Management Review, 24, 672-690.

LIPSEY, M. & POLLARD, J. 1989. Driving toward theory in program evaluation: more models to choose from. Evaluation and Program Planning, 12, 317-328.

LIPSKY, M. 1980. Street-level bureaucracy: Dilemmas of the individual in public services, New York, Russell Sage Foundation.

MARCHAL, B., VAN BELLE, S., DE BROUWERE, V. & WITTER, S. 2013. Studying complex interventions: reflections from the FEMHealth project on evaluating fee exemption policies in West Africa and Morocco. *BMC Health Services Research*.

MARCHAL, B., VAN BELLE, S., VAN OLMEN, J., HOERÉE, T. & KEGELS, G. 2012. Is realist evaluation keeping its promise? A literature review of methodological practice in health systems research. *Evaluation,* 18**,** 192-212.

MATLAND, R. 1995. Synthesizing the implementation literature: the ambiguity-conflict model of policy implementation. *Journal of Public Administration Research and Policy,* 5**,** 145-174.

MCLAUGHLIN, M. W. 1987. Learning from experience: lessons from policy implementation. *Educational Evaluation and Policy Analysis,* 9**,** 171-178.

MCPAKE, B., BRIKCI, N., COMETTO, G., SCHMIDT, A. & ARAUJO, E. 2011. Removing user fees: learning from international experience to support the process. *Health Policy Plan,* 26 Suppl 2**,** ii104-117.

MEESSEN, B., HERCOT, D., NOIRHOMME, M., RIDDE, V., TIBOUTI, A., TASHOBYA, C. K. & GILSON, L. 2011. Removing user fees in the health sector: a review of policy processes in six sub-Saharan African countries. *Health Policy Plan,* 26 Suppl 2**,** ii16-29.

MERTON, R. K. 1968. *Social theory and social structure,* New York, The Free Press.

MINGERS, J. 2000. The contribution of critical realism as an underpinning philosophy for OR/MS and systems. *Journal of the Operational Research Society,* 51**,** 1256-1270.

NIMPAGARITSE, M. & BERTONE, M. P. 2011. The sudden removal of user fees: the perspective of a frontline manager in Burundi. *Health Policy Plan,* 26 Suppl 2**,** ii63-71.

O'TOOLE, L. 2004. The theory-practice issue in policy implementation research. *Public Administration,* 82**,** 309-329.

O'TOOLE, L. & MONTJOY, R. 1984. Interorganzational policy implementation: a theoretical perspective. *Public Administration Review***,** 491-503.

OGRINC, G. & BATALDEN, P. 2009. Realist evaluation as a framework for the assessment of teaching about the improvement of care. *J Nurs Educ,* 48**,** 661-7.

PAWSON, R. 2013a. The science of evaluation: a realist manifesto, London, SAGE Publications.

PAWSON, R. 2013b. The Science of Evaluation. A Realist Manifesto., London, Sage.

PAWSON, R., GREENHALGH, T., HARVEY, G. & WALSHE, K. 2005. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy, 10 Suppl 1, 21-34.

PAWSON, R. & TILLEY, N. 1997. *Realistic Evaluation,* London, Sage.

RICHARD, F., ANTONY, M., WITTER, S., KELLEY, A., MEESSEN, B. & SIELEUNOU, I. 2013. Fee exemption for maternal care in Sub-Saharan Africa: a review of 11 countries and lessons for the region. *Global Health Governance,* 6**,**, 1-21.

RICHARD, F., WITTER, S. & DE BROUWERE, V. 2010. Innovative approaches to reducing financial barriers to obstetric care in low-income countries. *American Journal of Public Health,* 100**,** 1845-52.

RIDDE, V. 2011. What do we know in 2011 about the new free care policy in Mali? Montréal, Canada.

RIDDE, V. & DIARRA, A. 2009. A process evaluation of user fees abolition for pregnant women and children under five years in two districts in Niger (West Africa). *BMC Health Services Research,* 9**,** 89.

RIDDE, V., HADDAD, S., NIKIEMA, B., OUEDRAOGO, M., KAFANDO, Y. & BICABA, A. 2010. Low coverage but few inclusion errors in Burkina Faso: a community-based targeting approach to exempt the indigent from user fees. *BMC Public Health,* 10**,** 631.

RIDDE, V. & MORESTIN, F. 2011. A scoping review of the literature on the abolition of user fees in health care services in Africa. *Health Policy and Planning,* 26**,** 1-11.

ROBERT, E., RIDDE, V., MARCHAL, B. & FOURNIER, P. 2012. Protocol: a realist review of user fee exemption policies for health services in Africa. *BMJ Open,* 2**,** e000706.

ROSAS, S. 2005. Concept Mapping as a Technique for Program Theory Development: An Illustration Using Family Support Programs. American Journal of Evaluation, 26**,** 389-401.

RYCROFT-MALONE, J., FONTENLA, M., BICK, D. & SEERS, K. 2010. A realistic evaluation: the case of protocol-based care. *Implement Sci,* 5**,** 38.

TILLY, C. 2001. Mechanisms in Political Processes. *Annual Review of Political Science***,** 21-41.

TROCHIM, W. 1989. Pattern matching and program theory. *Evaluation and Program Planning,* 12**,** 355-366.

WALKER, L. & GILSON, L. 2004. 'We are bitter but we are satisfied': nurses as street-level bureaucrats in South Africa. *Soc Sci Med,* 59**,** 1251-61.

WITTER, S. 2009. Service- and population-based exemptions: are these the way forward for equity and efficiency in health financing in low income countries? . *In:* CHERNICHOVSKY, D. & HANSON, K. (eds.) *Advances in Health Economics and Health Services Research. Innovations in health systems finance in developing and transitional economies.* Emerald.

WITTER, S. & ADJEI, S. 2007. Start-stop funding, its causes and consequences: a case study of the delivery exemptions policy in Ghana. *The International Journal of Health Planning and Management,* 22**,** 133-43.

WITTER, S., ARHINFUL, D. K., KUSI, A. & ZAKARIAH-AKOTO, S. 2007a. The experience of Ghana in implementing a user fee exemption policy to provide free delivery care. *Reproductive Health Matters,* 15**,** 61-71.

WITTER, S., DIENG, T., MBENGUE, D., MOREIRA, I. & DE BROUWERE, V. 2010. The national free delivery and caesarean policy in Senegal: evaluating process and outcomes. *Health Policy and Planning,* 25**,** 384-92.

WITTER, S., GARSHONG, B. & RIDDE, V. 2013a. An exploratory study of the policy process and early implementation of free NHIS coverage for pregnant women in Ghana *International Journal for Equity in Health* 12**,** 1-11.

WITTER, S., KHADKA, S., NATH, H. & TIWARI, S. 2011. The national free delivery policy in Nepal: early evidence of its impact on health facilities. . *Health Policy and Planning,* 26**,** 84-91.

WITTER, S., KHALID, K., ABDEL-RAHMAN, M., HUSSEIN, R. & SAED, M. 2012. Removal of user fees for caesareans and under-fives in northern Sudan: a review of policy implementation and effectiveness. *nternational Journal of Health Planning and Management,* 28**,** e95-e120.

WITTER, S., KUSI, A. & AIKINS, M. 2007b. Effects of a delivery exemption scheme on workloads and incomes of health workers in Ghana. *Human Resources for Health* 5.

WITTER, S., MAKOUTODE, P., DOSSOU, J.-P., CUNDEN, N. & RICHARD, F. 2013b. Assessing the implementation of health financing policies – a proposed approach and discussion of its application to an emergency obstetric care exemption policy in Benin. *Global Health Action (under revision)*.

YANOW, D. 1987. Towards a policy culture approach to implementation. *Policy Studies Review,* 7**,** 101-115.

YIN, R. 2003. Case study research. Design and methods, London, Sage Publications.

YIN, R. 2009. Case study research. Design and methods, Los Angeles, Sage.