Did the learning agenda of the World Bank-administrated Health Results Innovation Trust Fund shape politicised evidence on performance-based financing? A documentary analysis

Lara Gautier¹
PhD Candidate at the University of Montreal, Montreal (Quebec), Canada
lara.gautier@umontreal.ca

Valéry Ridde²
Research Director at the French Institute for Research on Sustainable Development (IRD), Université Paris Descartes, Paris, France
valery.ridde@ird.fr

Abstract
The World Bank, co-funded by Norway and the United Kingdom, created and managed an innovative financing mechanism, the Health Results Innovation Trust Fund (HRITF), to support performance-based financing (PBF) reforms in low- and middle-income countries. From its inception in late 2007, until the closing of fundraising in 2017, it has carried out a wide range of activities related to experimenting PBF. In conjunction with the World Bank, which positioned itself as a “learning organisation”, donors have pushed the HRITF towards developing a specific learning agenda for documenting the policy impact of PBF. This learning agenda has been primarily based on impact evaluations of PBF pilot programmes. As a new body took over the HRITF’s portfolio (Global Financial Facility), a documentary analysis of this learning agenda is timely.

¹ Lara Gautier is also PhD Candidate at the Centre d’Etudes en Sciences Sociales sur les Mondes Africains, Américains et Asiatiques (UMR 245), IRD (French Institute For Research on Sustainable Development), Université Paris Sorbonne Cité, Paris, France.

² Valéry Ridde is also a researcher, Public Health Research Institute (IRSPUM), University of Montreal, Montreal (Quebec), Canada.
Building from public policy concepts that have been applied to social and health policy, and knowledge translation literature, we examine the learning agenda implemented by the HRITF over these 10 years. Our data includes documentation and publications (N=35) on HRITF and from the HRITF online platform. Results indicate that on several fronts, the HRITF shaped some form of politicised knowledge, notably in the ways country pilot grants were designed and evaluated. Some of its learning activities also provided opportunities for a transformative use of knowledge for World Bank staff as well as national implementers and policymakers. We also provide reflections about the HRITF's preferred approaches to produce knowledge and learn.

**Keywords:** Performance-based financing, literature review, documentary analysis, the World Bank, impact evaluation, politics of evidence.

“There is great interest from the global community to learn more about [health results-based financing, RBF] in a short timeframe. This interest stems not only from academics’ or practitioners’ interest, which is clearly prevalent, but also from a need to provide evidence on RBF in order to foster donor support and country investment, at a time when several countries have to address sustainability concerns.” (The World Bank, 2016a, p. 2)

### Introduction

The need to develop and use knowledge on the impact of healthcare system policies is a critical strand of work of major global health organisations. A large number of these policies addressed public healthcare systems financing in low – and middle – income countries (LMICs). Indeed, for the past 20 years, many healthcare financing strategies to improve supply, demand, and access to health services have been promoted and funded by international donors. These are piloted and scaled-up in spite of uncertainty as to their impact and effects on health systems. A report by the World Health Organization (World Health Organization, 2013) on achieving universal health coverage (UHC) – which, among all these strategies, has been winning donors’ attention – emphasises critical research gaps to be addressed.

Along with UHC, an approach mostly focusing on the supply-side of the healthcare financing equation was introduced: performance-based financing (PBF). PBF has been defined as a “policy innovation, whereby healthcare providers are, at least partially, funded on the basis of their performance” in attaining predefined healthcare targets (Gautier et al., 2018, p. 165). Pilot programmes of PBF have multiplied over the past fifteen years in LMICs. These programmes have been promoted, designed, funded, implemented, and evaluated by global actors (i.e., multilateral and bilateral donors, and non-governmental organisations) (Gautier; Ridde, 2017). In the mid-2000s, under the leadership of the Norwegian Agency for Development Cooperation (NORAD), the idea of a multi-donor trust fund emerged; and in December 2007, the Health Results Innovations Trust Fund (HRITF) was created. Administered by the World Bank, the HRITF’s core missions were to raise funding from donors, to offer technical assistance in countries and build
their institutional capacity to scale up and sustain PBF, and to produce and disseminate “evidence-based knowledge for a successful implementation of PBF” (RBF Health, 2016a). To date, HRITF has committed 385.6 million USD for 35 RBF programmes across 29 countries (Bhandari et al., 2017). This funding was matched to US$2.0 billion provided by the International Development Assistance fund (RBF Health, 2016a). This time period coincides with assessing the ending of the HRITF portfolio – whose fundraising function has been overtaken by the Global Financial Facility (Fernandes; Sridhar, 2017). As a policy idea, a wide range of individuals have challenged PBF (Paul et al., 2018) efficient and equitable approach to improving the performance of health systems in low-income and middle-income countries (LMICs) while its (individual and collective) promoters have sought to actively defend it (e.g., MayakaManitu, 2018; VanHeteren, 2018). This paper thus provides a timely analysis of the learning agenda of one of the major PBF players – the World Bank-managed HRITF.

In global health, analyses of such learning agendas have seldom questioned the use of evidence in policymaking (Lee; Goodman, 2002). Several global health researchers concur that “there has been limited attention on how financial resources used to gather evidence may have influenced its creation and presentation” (Hanefeld; Walt, 2015, p. 120). The HRITF case is interesting because some authors have uncovered potential biases in the evaluation of performance-based financing funded by the HRITF in LMICs (Barnes et al., 2014; Ireland; Paul; Dujardin, 2011; Turcotte-Tremblay et al., 2016). The portfolio of the HRITF may also provide valuable and genuine opportunities for learning in these countries. We consider policy diffusion, research on the politics of evidence, and the knowledge translation literature to investigate the politicisation of this global health institution’s learning agenda, which is administered by an international organisation (the World Bank) that refers to itself as “the knowledge bank” (Zack, 2003, p. 70). Our research investigates whether and how the learning agenda of the World Bank-administrated Health Results Innovation Trust Fund shape politicised evidence on performance-based financing. By reviewing the HRITF’s documentation and publications assessing the HRITF’s portfolio of activities, we aim to shed light on some of the pitfalls as well as opportunities induced by the implementation of the HRITF’s learning agenda.

Methods

i) Analytical framework

Drawing from policy diffusion, analysis of the politicisation of evidence, and the literature on knowledge translation, we build an analytical framework that guides our review of the HRITF’s learning agenda.

First, the body of policy diffusion literature can be useful to analyse PBF, because pilot programmes testing this policy innovation have flourished in low- and middle-income countries, sometimes leading to national policy adoption (Sina Health, 2017).
Many countries have learnt from one another. Yet some policy diffusion analysts tell us that learning policy experience from elsewhere raises an important risk of bias (Gilardi, 2010; Weyland, 2009). Investigating the diffusion of social policies, Weyland questions the ability of policy actors to “process the relevant information in a systematic, unbiased way” (Weyland, 2005, p. 263). Instead, they tend to “rely on cognitive heuristics that make it easier to select and digest an overabundance of information but that can also distort inferences significantly” (Weyland, 2005, p. 263). Therefore, policy actors engage in selection and digestion of information based on the cognitive frames available to them.

Cognitive frames have attracted a lot of attention from other strands of public policy literature (Béland; Cox, 2010; Cairney, 2016; Nature, [s.d.]; Parkhurst, 2016b). Drawing conceptual reflections from analysing the so-called movement of “evidence-based policymaking”, Parkhurst argues that on the contrary: evidence is political (Parkhurst, 2016b). Indeed, the way knowledge is conceived by policy actors matters because it reflects their belief system (Parkhurst, 2016a, p. 12). Yet policy actors tend not to consider these personal beliefs and motivations in their use and selection of knowledge. Parkhurst identifies two biases in the use of evidence in policymaking: a technical bias (i.e., political manipulation and cherry-picking of evidence) and an issue bias (i.e., in the creation of evidence and/or in the selection of the latter). As a form of technical bias, political manipulation happens when “scientific accuracy” is sacrificed when “policy decisions can determine the political or financial survival of involved actors” or where non-state actors “produce biased evidence in their interests” (Parkhurst, 2016a, p. 9). Cherry-picking of references towards sustaining scientific evidence making the case for a policy is another example of such bias. Issue bias may arise when policy actors are “unaware how their value systems, or their group identities, bias their understandings and interpretations of evidence” (Parkhurst, 2016a, p. 11). It follows that the legitimacy induced by the production and dissemination of evidence (i.e., publishing and presenting research findings through peer-reviewed articles, reports, and policy briefs (McSween-Cadieux et al., 2017)) may be partial without actors necessarily realising it. A common bias resulting from this process is the confirmation bias, which can be defined as: “the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand” (Nickerson, 1998, p. 175). Lastly, and most importantly, Parkhurst points to the (over)use of impact evaluations in policymaking arenas, which might illustrate another category of bias: that of “attribute substitution”. It is about “substituting the difficult questions of what to do to make society better with more straightforward questions of what interventions produce an effect” (Parkhurst, 2016a, p. 384). In other words, attribute substitution is choosing to pursue what can be measured (e.g., incidence rates or vaccination coverage), instead of looking at what might be more significant (e.g., social interactions or underlying structures) in societies. For instance, impact evaluations typically suffer from a representative bias, which derives from “representativeness heuristic” (Gilovich; Griffin, 2002), whereby assumptions are derived from “perception of similarity between a given situation and a prototypical one” (Parkhurst, 2016a, p. 383).
Lastly, the literature on knowledge translation provides relevant analytical lens to identify valuable opportunities that learning agendas may offer. McSween-Cadieux et al.’s empirically-drawn framework (Mc Sween-Cadieux et al., 2017, p. 8) identified three types of knowledge use by multiple policy actors (including international actors) in Burkina Faso. The first is related to politicisation (pervasive use), while the other two are “positive” knowledge uses: conceptual use, which enabled actual learning and skill development, and instrumental use, which fostered awareness-raising and change (Mc Sween-Cadieux et al., 2017).

From these theoretical underpinnings, it is possible to draw analytical categories relating to policy actors’ interaction with knowledge in global public health. Importantly, we take a different perspective from the literature that provides frameworks depicting how national policymakers interact with knowledge (e.g., Rodríguez et al., 2017). In this investigation, our study objects are not national policymakers but global actors pursuing policy diffusion. We identify four main categories of “knowledge processing”: selection, digestion, production, and dissemination of knowledge. The sequencing of these four categories may vary: for instance, dissemination of knowledge may represent the first step towards using knowledge. Parkhurst’s categories of biases are grouped into the category of “Persuasive use of knowledge”, while “positive” types of knowledge uses come under an umbrella named “Transformative use of knowledge”. Building from this, we draw an analytical framework that is applied to the case at hand (Table 1).

Table 1 – Overview of the analytical framework

<table>
<thead>
<tr>
<th>Knowledge processing categories</th>
<th>Persuasive use of knowledge</th>
<th>Transformative use of knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technical bias</td>
<td>Attribute substitution</td>
</tr>
<tr>
<td></td>
<td>Issue bias</td>
<td></td>
</tr>
<tr>
<td>Selection of knowledge</td>
<td>Cherry-picking of evidence</td>
<td>Confirmation bias</td>
</tr>
<tr>
<td>Digestion of knowledge</td>
<td>Cherry-picking of evidence</td>
<td>Confirmation bias</td>
</tr>
</tbody>
</table>
Production of knowledge

- Political manipulation of evidence
- Confirmation bias
- Representative bias; quantitative measurements of complex phenomena (e.g., impact evaluations)

Dissemination of knowledge

- N/A
- Selective dissemination of knowledge
- Tendency to overvalue results from studies evaluating the (quantitative) impact of policy innovations

Learning how to implement impact and process evaluations

Learning how to disseminate knowledge in meaningful and attractive ways

Adapted from: Weyland 2005; Parkhurst 2016; Mc Sween-Cadieux et al. 2017, and authors' own propositions

### ii) Data collection, extraction, and analysis

For this documentary analysis (Shaw; Elston; Abbott, 2004), we searched for two main types of data: internal resources of the HRITF, and external resources reporting on or analysing the HRITF portfolio of activities. For the former, we looked for online-available resources (including: manuals, reports, web stories, and PowerPoint presentations) extracted from the World Bank’s Results-Based Financing (RBF) Health web platform (http://www.rbfhealth.org). We also searched through two main World Bank databases: the World Bank’s “Open knowledge repository” database using the following search terms: “performance-based” or “results-based” or “impact evaluation”, as well as the World Bank Health and Nutrition “Documents & Reports” database using search terms: “performance-based” or “results-based”. We reviewed all the contents of the RBF web platform and screened for content specifically related to the learning agenda. Table 2 summarises the resources extracted.

**Table 2 – Titles and sources of documentation extracted from the RBF Health and from the World Bank’s website**

<table>
<thead>
<tr>
<th>Category</th>
<th>Title</th>
<th>Web source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webpage</td>
<td>History. The Health Results Innovation Trust Fund</td>
<td><a href="http://www.rbfhealth.org/mission-history">http://www.rbfhealth.org/mission-history</a></td>
</tr>
<tr>
<td>Webpage</td>
<td>Mission. The Health Results Innovation Trust Fund</td>
<td><a href="http://www.rbfhealth.org/mission">http://www.rbfhealth.org/mission</a></td>
</tr>
<tr>
<td>Webpage</td>
<td>Country Pilot grants; Knowledge and Learning grants [screening of all</td>
<td><a href="http://www.rbfhealth.org/projects">http://www.rbfhealth.org/projects</a></td>
</tr>
<tr>
<td></td>
<td>grants in these two databases]</td>
<td></td>
</tr>
<tr>
<td>Institutional</td>
<td>Learning Agenda for Results-Based Financing in the Health Sector: The</td>
<td><a href="https://www.rbfhealth.org/sites/rbf/files/The%20Health%20Results%20">https://www.rbfhealth.org/sites/rbf/files/The%20Health%20Results%20</a></td>
</tr>
<tr>
<td>strategy</td>
<td>Health Results Innovation Trust Fund Learning Strategy</td>
<td>Files/The%20Health%20Results%20Innovation%20Trust%20Fund%20Learning%20Strategy.pdf</td>
</tr>
<tr>
<td>Type</td>
<td>Description</td>
<td>Link</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Discussion paper</td>
<td>Qualitative Research to Enhance the Evaluation of Results-Based Financing Programmes: The Promise and the Reality</td>
<td><a href="http://www.rbfhealth.org/sites/rbf/files/Qualitative%20Research%20to%20Enhance%20the%20Evaluation%20of%20Results-Based%20Financing%20Programmes%20The%20Promise%20and%20the%20Reality_0.pdf">http://www.rbfhealth.org/sites/rbf/files/Qualitative%20Research%20to%20Enhance%20the%20Evaluation%20of%20Results-Based%20Financing%20Programmes%20The%20Promise%20and%20the%20Reality_0.pdf</a></td>
</tr>
<tr>
<td>Blog</td>
<td>Re-Imagining Results-Based Financing: Gearing up for the Future</td>
<td><a href="http://www.rbfhealth.org/blog/re-imagining-results-based-financing-gearing-future">http://www.rbfhealth.org/blog/re-imagining-results-based-financing-gearing-future</a></td>
</tr>
<tr>
<td>Blog</td>
<td>Scaling up and integrating your Results-Based Financing scheme: a progression in four phases</td>
<td><a href="http://www.rbfhealth.org/blog/scaling-and-integrating-your-results-based-financing-scheme-progression-four-phases">http://www.rbfhealth.org/blog/scaling-and-integrating-your-results-based-financing-scheme-progression-four-phases</a></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Blog</td>
<td>My learning experience: Results-Based Financing in Zimbabwe</td>
<td><a href="http://www.rbfhealth.org/blog/my-learning-experience-results-based-financing-zimbabwe">http://www.rbfhealth.org/blog/my-learning-experience-results-based-financing-zimbabwe</a></td>
</tr>
<tr>
<td>Blog</td>
<td>We just learned a whole lot more about RBF</td>
<td><a href="http://www.rbfhealth.org/blog/we-just-learned-whole-lot-more-about-rbf">http://www.rbfhealth.org/blog/we-just-learned-whole-lot-more-about-rbf</a></td>
</tr>
<tr>
<td>Blog</td>
<td>Results-Based Financing Writeshop: Improving implementers’ documentation and dissemination of experiences and lessons learnt</td>
<td><a href="http://www.rbfhealth.org/blog/results-based-financing-writeshop-improving-implementers%E2%80%99-documentation-and-dissemination">http://www.rbfhealth.org/blog/results-based-financing-writeshop-improving-implementers’-documentation-and-dissemination</a></td>
</tr>
<tr>
<td>Blog</td>
<td>Building Performance-Based Financing Knowledge Base Key to Successful Scale-Up of Programs</td>
<td><a href="http://www.rbfhealth.org/blog/building-performance-based-financing-knowledge-base-key-successful-scale-programs">http://www.rbfhealth.org/blog/building-performance-based-financing-knowledge-base-key-successful-scale-programs</a></td>
</tr>
<tr>
<td>Blog</td>
<td>Results-Based Financing: What Research Priorities Should We Address Over The Next Five Years?</td>
<td><a href="http://www.rbfhealth.org/blog/results-based-financing-what-research-priorities-should-we-address-over-next-five-years">http://www.rbfhealth.org/blog/results-based-financing-what-research-priorities-should-we-address-over-next-five-years</a></td>
</tr>
<tr>
<td>Blog</td>
<td>Welcome to the All Things RBF Blog!</td>
<td><a href="http://www.rbfhealth.org/blog/welcome-all-things-rbf-blog">http://www.rbfhealth.org/blog/welcome-all-things-rbf-blog</a></td>
</tr>
</tbody>
</table>
Thirty documents were selected: 15 blog posts, three annual/progress reports, four webpages, three toolkits, two discussion papers, two evaluation syntheses, and one institutional strategy. In addition to these 30 “internal” HRITF or World Bank resources, we looked for two types of documents investigating the HRITF’s portfolio. First, we selected relevant documentation on the HRITF from organisations linked to the World Bank (i.e., works ordered by the World Bank, its funders, and a major main think tank that influences the institution). Those include the following, respectively: the Norwegian Agency for Development Cooperation’s and United Kingdom’s Department for International Development’s official websites, online reports produced by the World Bank’s Independent Evaluation Group, and the Center for Global Development’s website. We identified four key references: an external evaluation from 2014 of the World Bank’s health financing strand of work (Schneider, 2014); the Center for Global Development’s special blog about the “HRITF at 10” (Bauhoff; Glassman, 2017); NORAD’s formal evaluation of the HRITF (Norwegian Agency for Development Cooperation, 2012); and the United Kingdom’s Department for International Development report on their involvement in results-based financing (Department for International Development, 2014).

Once we selected the data sources (35 references in total), we developed an Excel spreadsheet containing two types of entries corresponding to our analytical framework: knowledge processing stages (rows) and knowledge uses (columns). Subsequently, we extracted data related to the description of the HRITF’s learning agenda. We articulated the results by knowledge processing stages based on our analytical framework, shedding light on some of the pitfalls, but also on the opportunities induced by the implementation of the HRITF’s learning agenda. We discussed the results in light of 68 peer-reviewed articles addressing performance-based financing, whose search and selection process has been detailed in a separate paper (Gautier et al., 2018, p. 167).

Results

i) Overview of HRITF activities

Two countries – Norway and the United Kingdom – committed funds to the HRITF. Embracing the evidence-based policymaking injunction (Jones; Young, 2007), these two nations conditioned their financial provisions to the implementation of impact evaluations. Impact evaluations “assess the causal effects (impacts) attributable to an intervention by comparing the outcomes of interest (short, medium, or long term) with what would have happened without the program counterfactual” (Independent Evaluation Group, 2012). Since its inception, the HRITF thus had an explicit learning agenda primarily based on these impact evaluations of PBF pilot programmes (Schneider, 2014): “a well-funded impact evaluation portfolio underpins HRITF’s comprehensive learning agenda” (RBF Health, 2016b). In addition to impact evaluations, there were many learning activities developed to implement the HRITF learning agenda (Table 3).
**Table 3** – Activities developed to implement the HRTIF learning agenda in LMICs, 2008-2017

<table>
<thead>
<tr>
<th>Goals</th>
<th>Pilot programme implementation</th>
<th>Knowledge development</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Building technical capacity to implement PBF</strong></td>
<td>Funding training at country-level to implement pilot programmes</td>
<td>Funding international training of PBF implementers and/or high-level policymakers in LMICs (NB: training provided by international firms)</td>
</tr>
<tr>
<td><strong>Learning from PBF implementation elsewhere</strong></td>
<td>Contracting with international firms or consultants from other LMICs to provide technical assistance for successful implementation</td>
<td>Organising and funding study tours in “success story” countries (e.g., Rwanda, Burundi, Cameroon)</td>
</tr>
<tr>
<td><strong>Building country’s technical capacity to evaluate PBF</strong></td>
<td>N/A</td>
<td>Building capacity of countries to design, implement, and evaluate RBF programs</td>
</tr>
<tr>
<td><strong>Expand “evidence-based” knowledge on PBF</strong></td>
<td>N/A</td>
<td>Designing and implementing impact evaluations with quantitative components only (1st phase, from 2008 to 2012); with quantitative and qualitative components (2nd phase, from 2013 to now)</td>
</tr>
<tr>
<td><strong>Provide visibility to the HRTIF</strong></td>
<td>Creating and publicising online RBF courses and the “RBF game”, publishing annual reports and blogs sharing lessons from pilot experimentation</td>
<td>Publishing blogs, reports, and peer-reviewed articles that share impact evaluations results</td>
</tr>
<tr>
<td>Creating a community</td>
<td>Funding and organising international workshops for sharing lessons from pilot experimentation and impact evaluations results</td>
<td></td>
</tr>
</tbody>
</table>

**Source**: Authors’ compilation based on information available from the HRTIF website RBF Health: [www.rbfhealth.com](http://www.rbfhealth.com)

Thus, learning from PBF pilot programmes implementation was a critical aspect of the HRTIF’s learning agenda.
ii) A politicised selection of knowledge?

The second objective of the learning agenda was to “improv[e] the methods and measures used for assessing RBF (and determinants of its success)” (The World Bank, 2016a, p. 3). Measuring the effect of PBF was therefore critical. This evaluation involved first and foremost the selection of “measurable” indicators. The selection and setting of indicators in PBF represents a critical political moment: health providers get their rewards based on their commitment to achieving a number (quantity) of health services with technical quality (e.g., a qualified health staff should perform a certain number of antenatal consultations per month adequately recorded and followed-up). Since most PBF schemes initially targeted maternal and child health services, designers tended to refer to the same standardised list of indicators (e.g., Rusa et al., 2009). Besides, PBF scheme designers were already planning for impact evaluations and meta-analyses, therefore for comparability purposes, they thought it was best to harmonise lists of indicators. However, this approach resulted in attribute substitution issues. Indicators were often decided by World Bank people with their international partners (e.g., Vergeer et al., 2010), with inadequate consideration for contextual features (Paul; Sossouhounto; Eclou, 2014). At times, indicators did not match the human resources configuration of health facilities in LMICs. For instance, in Mali (The World Bank, 2017a), it is debateable to only reward deliveries assisted by qualified health workers, while birth attendants (“matrones”) are in most rural facilities the only staff available to perform such task.

Besides, despite the original idea that HRITF would support results-based financing broadly, i.e. encompassing both supply-side and demand-side incentives, emphasis has been almost exclusively set on the provider. Thus, apart from a few pilot experiences (in Benin, Burkina Faso and Cameroon, for instance), indicators have eluded the inclusion of the worst-off in PBF pilot schemes (De Allegri et al., 2018; Ridde et al., 2018a, 2018b). In response to this criticism, the HRITF portfolio has paid an increased attention to equity (The World Bank, 2015). More recent pilot schemes (e.g., in the Republic of Congo and in Central African Republic) have more systematically included the targeting of vulnerable people (“indigents”). Such schemes comprise an indicator whereby providers are rewarded more if they attend indigents (who do not pay or only pay reduced user fees at the point of service).

iii) Digesting types of knowledge: unravelling the HRITF’s learning-by-doing approach

The HRITF’s work represents one of the most salient examples of the World Bank deliberately and expressively adopting a learning-by-doing approach in health. Besides the development of impact evaluations of HRITF-funded projects (included in CPGs), it also relies on developing non-scientific knowledge through the so-called “knowledge and learning grants” (KLGs) which include funding for in-country and international training and study tours to featured countries (McCune, 2014).
This learning-by-doing vision guiding the learning agenda of the HRITF may yield transformative uses of evidence. A critical illustration of such potential transformation is the explicit goal to facilitate technical peer-to-peer dialogue across countries through funding and organising study tours (The World Bank, 2009). However, there might be concerns surrounding the “copypasting” approach that emerge: not all the activities implemented by “success story” countries like Rwanda may be adapted to other countries. This type of cherry-picking (learning only from the “best model”) leads to overconfidence in a standard model that is bound to mislead country implementers, thereby inducing cognitive shortcuts. From 2015 however, the HRITF has encouraged national delegations to visit other countries that have adopted different PBF models, such as Argentina (McCune, 2015), Burundi (Idrissi; Driss, 2015), or Zimbabwe (Socorro, 2016).

The HRITF online platform (http://www.rbfhealth.org) is a resource repository that serves as the main knowledge management tool operationalising the “learning-by-doing” vision. Notably, it provides online resources for skills development, including trainings, webinars, and the “RBF game”. Importantly, it includes a blog initially named “All Things RBF”, which “shares stories from practitioners and implementers around the world, including project experiences and personal perspectives from subject matter experts on a range of RBF topics” (Vledder, 2012). The blog started off with overly enthusiastic language (one of the first blog posts was entitled: “Results-Based Financing: A Proven Model for Better Maternal and Child Health” (Vledder, 2013)) which appeared to convey substantial confirmation biases. Conducting a discourse analysis of the content of this blog until July 2014, some authors argued that “[n]one of the 38 blog entries […] were overtly critical or specific about potential limitations of PBF” (Barnes et al., 2014, p. 25). Gradually however, we found that the content of blog posts became more reflective and, more recently, even critical. As an illustration, the first paragraph of a blog entry (by Loevinsohn; Nair, 2017) includes the following sentences: “After more than 8 years of implementing RBF in the health sector, this narrow focus on incentives as the sole driving force for results seems too narrow. Although RBF provides a common approach to thinking about improving the quality, delivery, and coverage of essential services, it is not ‘one size-fits-all’ by any means”. Uncertainty as to the effectiveness of PBF (Das; Gopalan; Chandramohan, 2016; Ogundeji; Bland; Sheldon, 2016;Witter et al., 2012), raises a few concerns for policymakers’ digestion of information on PBF. While World Bank staff acknowledged that evidence was mixed (Kandpal, 2016), there may have been some discrepancy in the diffusion of this information (Paul et al., 2018). Considering that policymakers in LMICs usually have to overcome technical and cognitive barriers to access to research (Hyder et al., 2011), governments with no or little experience with PBF might not have had access to such mixed evidence. This lack of informed decision making prior to engaging in pilot schemes may have led to a confirmation bias, where by government representatives expected PBF to deliver on its promises – based on the moral and financial authority of the World Bank.
iv) Producing multiple types of knowledge?

The HRITF learning agenda entailed the production of multiple types of knowledge. For instance, the RBF Health website provides a number of resources (e.g., reports, evidence syntheses, toolkits, etc.) documenting the impact of PBF. But the core learning component of the HRITF’s learning agenda lied in producing impact evaluations. According to its last progress report, 33 impact evaluations were funded by HRITF across 28 countries (Health Results Innovation Trust Fund, 2016). In general, the multiple roles that HRITF, managed by the World Bank, undertook in country pilot grants (CPGs) can be problematic. Indeed, when a Task Team Leader in charge of a pilot scheme (responsible for taking decisions on disbursements, and setting the rules of collaboration for PBF pilot implementation) is also involved in the design and implementation of an impact evaluation of the same program, there is a high risk of confirmation bias. Although one of the core principles of PBF is the separation of functions, few researchers have looked into this issue (Turcotte-Tremblay et al., Forthcoming).

One could argue that concentrating on the sophisticity of impact evaluation designs and on quantitatively measurable data, has been pursued at the expense of policy relevance, leading to the phenomena of attribute substitution. Indeed, attention to operational complexities or intermediary mechanisms may be more useful to policymakers in LMICs. First, expectations about the design of complex (i.e., multi-arm and randomized) models of impact evaluations sometimes differed between the funder and governments’ representatives. Some would argue that multi-arm models (applied in Burkina Faso, Cameroon, and Zambia) are “incredibly useful” means for understanding “the most cost-effective strategy to increase coverage and quality, and to establish attributable impact outcomes” (Bauhoff; Glassman, 2017, p. 2). Yet their implementation involved lengthy, heavy, and sometimes even controversial processes (Gautier et al., 2018). In Burkina Faso for instance, there was a discrepancy between what World Bank/HRITF staff wanted to undertake for the impact evaluation, and what local actors actually wished for. The design included four arms to test the impact of PBF alone versus together with other interventions (including health insurance). This testing involved developing a complicated process of randomization – in over 500 facilities – that was criticised by local actors (Ridde et al., 2017). Given the institutional orientation towards a universal health insurance scheme (ibid.), it might especially be counter-intuitive to ensure that one of the control groups would not implement insurance. Besides, the national user-fee exemption policy for children under five and pregnant women implemented in 2016 (Zombré; De Allegri; Ridde, 2017a) may undermine the current impact evaluation: it will be difficult to disentangle the effects of this policy from the effects of PBF. In Cameroon as well, there were complaints about the randomization process (also including four arms (De Walque; Robyn; Sorgho, 2013)): several health providers expressed a feeling of injustice (RBF Health, 2014). Second, the support provided by the HRITF included skills development in evidence-
informed policymaking: “capacity building among country teams for implementing impact evaluations” (Elridge; Tekolste, 2016, p. 3). Considering the frustrations expressed by some country representatives, assessing the effects of this approach is significant. Most importantly, making sure county teams understand that policy relevance is more critical than developing sophisticated impact evaluations designs is essential.

To date, evaluating country pilot grants (CPGs) through quantitative methods have remained the fund’s priority: very few qualitative components have been included in the evaluative design. A qualitative component was added in only 13 out of these 28 countries (Cataldo; Kielmann, 2016). More importantly, one could argue that process evaluations may be missing from the picture (except for Cameroon where a process evaluation was recently published (De Allegri et al., 2018)), even though they are extremely useful to unpack the “how” – the circumstances in and conditions for which an intervention can deliver expected results. The reasons why HRITF decided not to fund more process evaluations – which could have brought critical and complementing elements to impact evaluations - may lie in the preference for quantitative measurements, which are considered “more straightforward”, and above all, more controllable than qualitative inquiries.

Another issue arises with the impact evaluation strand of work: the World Bank (which coordinates the HRITF) appears to have taken a double standards approach to producing scientific knowledge on PBF. Several PBF pilot projects have been funded by the World Bank outside of the HRITF facility (e.g., in Mali and Niger (The World Bank, 2017a, 2017b)): these do not involve scientific evaluations. Generally speaking, the lack of scientific evidence coming from these countries’ experience with PBF is problematic because it creates inconsistency. On the one hand, in HRITF-funded CPGs, huge amounts of money have been spent on impact evaluations [The average cost of an impact evaluation done by the World Bank is close to 1 million USD. In Guinea, an impact evaluation of performance-based incentives in education amounted to over 2 million USD (Gertler et al., 2016, p. 217).], and on the other hand, non-HRITF funded pilots have not involved any instrument to produce what is referred to as “robust evidence”. It may notably create representative biases. For example, in Mali, only internal reports have been produced on the two pilot schemes implemented in the Koulikoro region (The World Bank, 2017a; Toonen et al., 2014). These pointed to positive outcomes of PBF in that region without controlling for confounding factors. However a concurrent independent investigation using a quasi-experimental design showed no effect on health utilisation (Zombré; De Allegri; Ridde, 2017b) of the first PBF experience. In addition, this research was not cited in the World Bank report that was written afterwards, which illustrates an example of cherry-picking of evidence (The World Bank, 2017a).

v) Disseminating the knowledge

In coherence with the evidence-based policymaking discourse, the underlying idea of the large impact evaluation strand of work, was that the ensuing dissemination of positive results would prompt policymaking (in favour of national scale-up, for instance).
Yet, it remains unclear as to how much actual evidence-based learning recipient countries’ representatives was achieved”. There are indications that this evidence-based approach to policymaking has not been very effective. In several countries (like Burkina Faso, Benin, or Argentina), “decisions were made to scale up regardless of weak, inconclusive, or incomplete pilot results” (Schneider, 2014, p. 55). One limitation for evidence-informed policymaking includes the fact that data collected by the World Bank through baseline, midline, and endline surveys for impact evaluations undergo an embargo of about two years. The idea of an “open” knowledge bank (Kiendrébéogo, 2014) providing readily available data on PBF pilot schemes is therefore misleading: if policymakers need to wait for years before independent researchers can use this data, and disseminate their own evidence, it may be more difficult for them to make informed and balanced decisions.

Besides the publication of peer-reviewed studies reporting on impact evaluation results, the learning strategy included multiple knowledge dissemination activities. HRITF-funded projects included the organisation of yearly gatherings held in multiple places across the world (e.g., in Thailand in 2011, Turkey in 2012) to disseminate early findings from impact evaluations to national actors implementing PBF pilot schemes funded by the HRITF. Soon, it became clear that these international workshops would also represent relevant fora for sharing decision makers and practitioners’ lessons learnt (e.g., in Argentina in 2015, and Zimbabwe in 2016) (Kiendrébéogo, 2014). This opportunity contributed to the development of a community of PBF practitioners (Barnes; Brown; Harman, 2015). The contents of these workshops were often restricted to technical matters (e.g., McCune, 2015). However, with time genuine exchanges of lessons learnt could be shared on these occasions, including on the (not only technical) challenges posed by implementing PBF. This happened, in particular, at the last workshop held in Zimbabwe (The World Bank, 2016b). At the end of the day, actors mostly benefiting from listening to contextualised policy-related challenges (e.g., Jansen; Toonen, 2016) are likely the members of the World Bank Research Group themselves. The shift, starting from late 2012, towards including qualitative components to impact evaluations (Hasan, 2012) might have been the most salient outcome of such exchanges, thereby demonstrating adaptation on the researchers’ side – in order to fill research gaps.

In many ways, the learning-by-doing approach may have pushed the HRITF along the path of a “learning organisation” (Akhnif et al., 2017), which includes and values not only scientific evidence but also practice-based expertise and participatory co-constructed knowledge. There are bold indications that the HRITF developed strong skills in meaningful and attractive ways. The recent set up of “RBF writeshops” indicates the HRITF’s will to value and promote practitioners’ lessons learnt, by teaching them how to document these lessons, and coaching them in developing articles on the topic of their choice (Josephson, 2017). Future research should look into the how and why the learning-by-doing approach is useful to enhance the World Bank’s credibility and legitimacy in healthcare financing. Yet, ensuring constant questioning and debating of the knowledge that is produced before it gets disseminated (Health Results Innovation Trust Fund, 2016) remains critical in order to
avoid any accusation of selective dissemination of knowledge, based on the observed World Bank’s mingling of advocacy and knowledge production roles.

**Discussion**

**A politicised learning agenda?**

The most frequently cited illustrations of a politicised selection, digestion, production, and dissemination of knowledge are attribute substitution and issue biases.

First, in designing PBF schemes, risks of attribute substitution manifested in the inadequate consideration for contextual features in selecting the policy’s indicators. Several papers pointed to a lack of ownership by decision makers (Chimhutu et al., 2015; Gautier; Ridde, 2017). Moreover, the direct beneficiaries of PBF, i.e. health providers at the individual or collective level, were rarely consulted to provide input to the design of PBF schemes in their country/area. In Benin, such lack of contextual consideration has caused incomplete adherence by implementers of the PBF scheme (Paul; Sossouhounto; Eclou, 2014). In Nigeria, an independent study has shown that scepticism about the adequacy of healthcare workers’ assessment tools led to incomplete adherence (Ogundeji; Bland; Sheldon, 2016). Therefore, there is a need for enhanced policy relevance in countries where PBF is tested. There needs to be a more adaptive, contextualised type of assistance for designing PBF pilot schemes. At this stage, letting country representatives decide what would fit their epidemiological profile and health priorities/health system general planning would be a first step, and would comply with the Paris Declaration on aid effectiveness (OECD, 2005). Trying to integrate or build from previous experience(s) of PBF, even if other donors were involved, would create the conditions for more constructive collaborations with government and other international partners.

Efforts to widen the scope of indicators were not always conclusive. For instance, as indicated in the results, the HRITF portfolio started to introduce more indicators to link incentives to equity performance. Yet, as a recent paper on the Cameroon experience shows, providing higher rewards for the care of the worst-off “does not seem to be enough to effectively reach disadvantaged populations and increase care among the very poor” (De Allegri et al., 2018, p. 7). Providers get individual and collective financial rewards, while patients (including vulnerable ones) still have to pay to access health services.

The cherry-picking of the “right model” to reproduce in other LMICs also raises concerns for policymakers’ digestion of knowledge. Indeed, getting primary “inspiration” from the Rwandan model (Meessen; Soucat; Sekabaraga, 2011) through funding and organising study tours and workshops in that country is problematic. In Rwanda, the “mutuelles” model has equally been copy-pasted in many other African countries, often without paying adequate attention to the peculiarities of the Rwanda contexts that made this “success” possible (Chemouni, 2018). The Rwandan experience, which served as “proof of concept” thanks to the publication of a paper in the renowned Lancet
Did the learning agenda of the World Bank-administered Health Results Innovation Trust Fund shape politicised evidence on performance-based financing? A documentary analysis

Lara Gautier (University of Montreal); Valéry Ridde (Université Paris Descartes).


(Basinga et al., 2011), also raises a few questions. The context of Rwanda – recovering from traumatising events under the leadership of a vocal President – is arguably very specific, with a peculiar political settlement, a complete legitimation project, and specific political ideas (Chemouni, 2018). Besides, the conclusion of the Lancet paper were partly questioned by subsequent papers (Ngo; Sherry; Bauhoff, 2016; Skiles et al., 2015) using the same Demographic Health Survey data.

Most importantly, this documentary analysis shows that the World Bank staff was involved in the promotion, design, funding, implementation (through technical assistance), and evaluation. This involvement may have created a conflict of interest (Gautier; Ridde, 2017; Turcotte-Tremblay et al., 2016). Even though impact evaluations were designed and analysed by a different World Bank group (the Research group, whereas the Health Nutrition Population unit was in charge of pilot implementation), the same institution engaged with each of the stages (i.e. from designing pilot schemes to results dissemination). While some research may have been conducted by external parties, the relationship involved controlled research findings disclosure. This issue starts to get more research attention: authors have shown the possible pitfalls associated with such relationships (Doherty et al., 2018). As mentioned in the results section, this connection raises major risks of confirmation biases in the production of knowledge.

The initial focus on producing quantitative impact evaluations also requires commentary. The emergence of PBF coincided with the renaissance of impact evaluations using quasi-experimental research designs, which had lost credence in the 1980s (Shadish; Cook, 2009). Unlike PBF, community-based insurance schemes mainly expanded in the 1990s: there were very few impact evaluations with quasi-experimental designs published at the time, compared to recent years (Ekman, 2004; Raza et al., 2016; Spaan et al., 2012). Research trends in the 2000 decade, shaped by influential economists such as Esther Duflo, re-emphasised the value and relevance of undertaking impact evaluations to assess the effectiveness of development interventions (Duflo; Glennerster; Kremer, 2007). The main purpose of these evaluations was to draw scalable lessons from these “rigorous” impact evaluation results. Recently, there has been a lot of criticism of this movement, particularly from Nobel-prized economists (Bédécarrats; Guérin; Roubaud, 2017; Deaton; Cartwright, 2018). Critics pointed to the many pitfalls (frequently featuring representative biases) of generalising conclusions drawn from quasi-experimental evaluations, which design actually entails a necessary narrow scope of measureable effects, and which do not account for the many contexts in which the intervention takes place.

As observed in the results section, the content of HRITF’s knowledge dissemination workshops were – at least initially – restricted to technical matters (e.g., McCune, 2015). This constriction tended to portray the World Bank as a mere “passive facilitator of learning in global health policy” (Barnes; Brown; Harman, 2015, p. 88). However, such passivity might hide a feature of Ferguson’s “antipolitics machine” (Ferguson, 1990), whereby international organisations pursue the constant depoliticisation of complex social issues. Applying this argument to the case at hand would mean for the World Bank – and other
international actors promoting PBF – to be pursuing a strategic depoliticising of PBF-related issues, by reframing those as exclusively technical (Barnes et al., 2014, p. 23). For instance, Barnes and colleagues observe that for promoters (e.g., in Meessen; Soucat; Sekabaraga, 2011) “PBF schemes form contractual relationships rather than hierarchical ones” (Barnes et al., 2014, p. 23). However, issues of power and hierarchies in health facilities in LMICs may not be dealt with simply by issuing contracts between parties.

In general, what the HRITF’s learning strategy reveal is that donors offered to fund the HRITF for widening the body of knowledge on PBF, without actually knowing which way it could go (Bollinger; Kruk, 2016). This outcome is coherent with the results of two Cochrane systematic reviews (Witter et al., 2012; Wiysonge et al., 2017). The 2012 review notably pointed to important concerns over the imputability of observed effects to PBF as described in the studies, i.e., whether the PBF intervention was independent of other changes occurring at the same time as the intervention, such as other policies in place or seasonal features (Witter et al., 2012).

Opportunities for transformative learning

There were many indications that the HRITF did engage in transformative use of knowledge. The inclusion of qualitative components in impact evaluation designs, as well as the implementation of writeshops are salient examples of an authentic “learning-by-doing” approach, whereby World Bank researchers acknowledge the complex phenomena generated by PBF, and whereby PBF implementers acquire skills. Besides, the genuine peer-to-peer knowledge and practice-based exchanges created on the occasions of dissemination workshops and/or study tours equally represent positive opportunities for contextual adaptation of PBF schemes in LMICs. Pursuing the constant debating of knowledge is critical to avoid any accusation of selective dissemination of knowledge, based on the observed World Bank’s mingling of advocacy and knowledge production roles.

Building on these opportunities, a few policy recommendations can improve the relevance and appropriateness of the PBF portfolio. Primarily, the GFF could learn from the several pitfalls identified in this paper and in the numerous process evaluations published by researchers independently from HRITF-funded pilot programmes (e.g., Chimhutu; Lindkvist; Lange, 2014; Feldacker et al., 2017; Ridde et al., 2017; Turcotte-Tremblay et al., 2017), so as to initiate a reflection on how to re-orient the ways schemes are designed, implemented, and evaluated in LMICs. In particular, the GFF (who has taken over the RBF portfolio) should make sure that the lack of evidence-informed policymaking, which has prevailed for now, does not perpetuate. The GFF might also capitalise on the numerous achievements of the HRITF as a participatory “learning organisation”. Lastly, the possible conflict between the “advocacy role” and “knowledge-production role” of the World Bank is a critical aspect that should be further investigated with interview data.
Conclusion

This paper assesses the learning agenda of the HRITF using the available documentation on the subject. It concludes with a nuanced portrayal of the various activities undertaken by the HRITF and the World Bank to expand knowledge on the policy impacts of PBF. In the ways country pilot grants were designed and evaluated, the HRITF shaped some form of politicised knowledge. Several learning activities also provided opportunities for a transformative use of knowledge for World Bank staff as well as national implementers and policymakers.

This piece took an interdisciplinary approach (using public policy and knowledge transfer literatures) to perform a documentary analysis. The various dimensions of the analytical framework proved useful to make sense and organise the rich information extracted from the grey literature. However, the fact that we primarily relied on documents made it difficult to sustain arguments with specific examples and evidence for illustrating some of the analytical dimensions (e.g., those in the transformative use of knowledge category). This difficulty is an important limitation of this study. An in-depth qualitative investigation including participant observation of World Bank processes and activities would provide more depth in the analysis and unpack implications for future knowledge management strategies at the World Bank. The institution would advance towards becoming a transformative learning organisation should it accept to host such research approach.

References


Did the learning agenda of the World Bank-administrated Health Results Innovation Trust Fund shape politicised evidence on performance-based financing? A documentary analysis

Lara Gautier (University of Montreal); Valéry Ridde (Université Paris Descartes).


A agenda de aprendizado do Fundo Fiduciário de Inovação em Resultados para a Saúde, administrado pelo Banco Mundial, moldou evidências politizadas sobre financiamento baseado em desempenho? Uma análise documental

Resumo
O Banco Mundial, cofinanciado pela Noruega e Reino Unido, criou e gerenciou um mecanismo inovador de financiamento, o Fundo Fiduciário para Inovação dos Resultados na Saúde (HRITF), para apoiar reformas de financiamento baseado em desempenho (FBD) em países de renda baixa e média. Desde o seu início no final de 2007, até o fechamento da captação de recursos em 2017, realizou uma ampla gama de atividades relacionadas à experimentação do FBD. Em conjunto com o Banco Mundial, que se posicionou como uma “organização que aprende”, os doadores levaram o HRITF a desenvolver uma agenda de aprendizagem específica para documentar o impacto político (policy) do FBD. Esta agenda de aprendizado tem sido baseada principalmente em avaliações de impacto de programas pilotos de FBD. Considerando-se que um novo órgão assumiu a carteira do HRITF (Global Financial Facility), uma análise documental dessa agenda de aprendizado mostra-se oportuna. Tendo por base conceitos de políticas públicas aplicados à política social e de saúde, e a literatura de tradução de conhecimento, examinamos a agenda de aprendizagem implementada pelo HRITF ao longo desses 10 anos. Nosso dados incluem documentação e publicações (N = 35) sobre o HRITF e da plataforma on-line do HRITF. Os resultados indicam que, em várias frentes, o HRITF moldou alguma forma de conhecimento politizado, principalmente no modo como as doações para projetos pilotos foram desenhadas e avaliadas. Algumas de suas atividades de aprendizado também proporcionaram oportunidades de uso transformador do conhecimento pelos funcionários do Banco Mundial, bem como de implementadores e formuladores de políticas nacionais. Também apresentamos reflexões sobre as abordagens preferidas pelo HRITF na produção do conhecimento e aprendizado.

Palavras-chave: Financiamento baseado em desempenho, revisão da literatura, análise documental, Banco Mundial, avaliação de impacto, política de evidência.

¿La agenda de aprendizaje del Fondo Fiduciario para la Innovación en Resultados de la Salud administrado por el Banco Mundial formuló evidencia politizada sobre el financiamiento basado en el desempeño? Un análisis documental

Resumen
El Banco Mundial, cofinanciado por Noruega y el Reino Unido, creó y administró un mecanismo de financiación innovador, el Fondo Fiduciario para la Innovación en los Resultados de la Salud (HRITF), para respaldar reformas de financiamiento basada en el desempeño (FBD) en países de ingresos bajos y medianos. Desde su inicio a fines de 2007, hasta el cierre de la recaudación de fondos en 2017, ha llevado a cabo una amplia gama de actividades relacionadas con la experimentación de FBD. Junto con el Banco Mundial, que se posicionó como una «organización de aprendizaje», los donantes han impulsado el HRITF hacia el desarrollo de una agenda de aprendizaje específica para documentar el impacto de la política del FBD. Esta agenda de aprendizaje se ha basado principalmente en evaluaciones de impacto de programas piloto de FBD. A medida que un nuevo organismo se hizo cargo de la cartera de HRITF (Global Financial Facility), una análise documental dessa agenda de aprendizado mostra-se oportuna.
Did the learning agenda of the World Bank-administrated Health Results Innovation Trust Fund shape politicised evidence on performance-based financing? A documentary analysis
Lara Gautier (University of Montreal); Valéry Ridde (Université Paris Descartes).

Financial Facility), un análisis documental de esta agenda de aprendizaje es oportuno. A partir de los conceptos de política pública que se han aplicado a las políticas sociales y de salud, y la literatura de traducción de conocimiento, examinamos la agenda de aprendizaje implementada por HRITF durante estos 10 años. Nuestros datos incluyen documentación y publicaciones (N = 35) en HRITF y de la plataforma en línea HRITF. Los resultados indican que en varios frentes, el HRITF dio forma a algún tipo de conocimiento politizado, especialmente en la forma en que se diseñaron y evaluaron los subsidios de los países. Algunas de sus actividades de aprendizaje también brindaron oportunidades para un uso transformador del conocimiento para el personal del Banco Mundial, así como para los implementadores nacionales y los formuladores de políticas. También proporcionamos reflexiones sobre los enfoques preferidos de HRITF para producir conocimiento y aprender.

Palabras clave: Financiamiento basado en el desempeño, revisión de literatura, análisis documental, el Banco Mundial, evaluación de impacto, política de evidencia.

Acknowledgements: The authors would like to express their gratitude to Dr Benjamin Chemouni, who graciously offered to comment on an earlier version of this manuscript.

Received: 14/Apr/2018
Approved: 13/Jul/2018