Centre of Excellence for Development Impact and Learning

February 2022

Research project paper 1

How do we shape and navigate pathways to social accountability scale? Introducing a middle-level Theory of Change

Florencia Guerzovich, Tom Aston, Brian Levy, Paula Chies Schommer, Rebecca Haines, Sue Cant, Grazielli Faria Zimmer Santos

LIDC London International Development Centre

UKaid from the British people

Oxford Policy Management

About CEDIL

The Centre of Excellence for Development Impact and Learning (CEDIL) is an academic consortium supported by the UK government through UK Aid. The mission of the Centre is to test innovative methodologies in evaluation and evidence synthesis and promote evidence-informed development. CEDIL-supported projects fall into three programmes of work: evaluating complex interventions, enhancing evidence transferability, and increasing evidence use.

CEDIL research project paper

CEDIL research project papers are the research outputs of the projects supported by CEDIL. They include impact evaluations, systematic reviews, evidence and gap maps and exploratory papers on research methods and approaches.

About this research project paper

This paper was submitted to CEDIL by the Social Accountability Learning and Theory Project S.156 team.

Please direct any comments or queries to the corresponding author, Florencia Guerzovich at florcig@gmail.com

Suggested citation: Guerzovich, F., Aston, T., Levy, B., Chies Schommer, P., Haines, R., Cant, S., Faria Zimmer Santos, G. (2022) 'How do we shape and navigate pathways to social accountability scale? Introducing a middle-level Theory of Change', CEDIL Research Project Paper 1. Centre of Excellence for Development Impact and Learning (CEDIL), London and Oxford. Available from: https://doi.org/10.51744/CRPP1

This project was funded by the Centre of Excellence for Development Impact and Learning (CEDIL), supported by UK aid from the UK Government. The views expressed in this research project paper do not necessarily reflect the UK Government's official policies or CEDIL.

Acknowledgements

We are grateful to all those who supported this work. They include the World Vision team, including in the London, Afghanistan, Cambodia, DRC, Indonesia and beyond. Special thanks to Besi Mpepo, Isaiah Nzima, Richard Dove, Diane Papelleras, Marvin Patnugot, and Holta Trandafili. Jeff Thindwa, Maria Poli, Ann-Sofie Jespersen, and other members of the GPSA team and Steering Committee. Thanks to CEDIL and DFID/FCDO for partial funding, comments, and a community to discuss these issues. Howard White, Chris Roche and Marcella Vigneri, participants in two CEDIL partners Workshops and two anonymous reviewers. The OPM team provided support for grant management. The Ethics team and colleagues at ESAG/UDESC provided useful guidance. We are also grateful to David Jacobstein, Helene Grandvoinnet who provided comments to the initial proposal and/or earlier versions of this paper. Many other colleagues provided input to sub-products or background research for this paper we are grateful to each one of them Participants at the GPSA Global Partners Forums of 2019 and 2021 a side-consultation session we conducted in 2019, members of the WHO Community of Practice on Social Accountability and others who provided feedback. We are thankful to Johanna Wheeler for superb edition.

Abstract

In the last two decades, the uptake of social accountability interventions has grown exponentially. While evidence of social accountability's contribution to health outcomes is mixed, there is significant scale up of these interventions in the sector. This apparent contradiction is the point of departure for the development of middle-level theory on potential pathways to scale up in the health sector. The paper argues that practitioners can pursue at least three pathways to scale: the replication of best practice, through leveraging the countervailing power of resistance, and seeking resonance with existing public sector efforts. These figure in a nested theory of change framework. Each of these sub-theories may be better bets to scale in different contexts or at different moments of time in these contexts. Each pathway has a different normative point of departure and perspective on how scale up is achieved. Change may happen chiefly in response to: (1) new information and rigorous evidence (best practice), (2) civic pressure from the outside, paired with disruptors on the inside (resistance), or (3) deliberation, compromise, and collective action (resonance). Each pathway places different emphasis on the dividends derived from conflict and on the promise of social learning to resolve collective action problems. In unpacking these pathways through process tracing and comparative analysis, this paper seeks to overcome a zero-sum battle to better inform practitioner and policy-maker bets in the health sector, towards different pathways that work in different settings and interrogating the transferability of each pathway to scale (Masset and White, 2019). For donors, different combinations of these sub-theories may make up a more promising portfolio for advancing scaling up social accountability to deliver specific service delivery goals in concrete contexts and time periods on their own or by pursuing coordination with others over time.

Contents

I. Policy relevance: Why focus on scaling up social accountability?

Social accountability is a process that enables the inclusive participation and collective action of citizens and civil society organizations in public policymaking and implementation, so that state and service providers are responsive to citizens' needs and held accountable (Guerzovich and Poli, 2020). The practice of social accountability has steadily grown in the last two decades, in a number of sectors including in the health sector. While the evidence regarding the contribution of citizen-led social accountability to service delivery is mixed,[1] social accountability has become prevalent in global policy declarations, and national policies and programs. In this study, we focus on social accountability interventions initiated by non-governmental organized groups (rather than state entities). These interventions are initiated by these Civil Society Organizations (CSOs), including in the delivery of services[2].

The goal of the SALT project is to develop a middle-level theory that describes types of pathways social accountability practitioners can pursue to scale up, as well as the conditions under which each path might be more promising. Our work seeks to inform the decisions of international development partners who largely fund CSOs and CBOs that are implementing organizations for social accountability interventions, as well as those implementing organizations.

Major development institutions place high hopes on social accountability. For example, the UK Department for International Development's (DFID) *Governance for Growth, Stability and Inclusive Development* position paper noted that accountable institutions can promote more equitable, sustained economic growth and contribute to prosperity more generally. The paper also states that social accountability processes can help deliver better services (DFID, 2019: 8). This apparent contradiction – between the mixed evidence on results and the high expectations for social accountability - is the starting point for the project (*Scaling Social Accountability for Health: Leveraging Public Policies and Programs*, or SALT) and informs this paper.

[1] For a discussion of mixed evidence and why the evidence will continue to be mixed, see Cant and Sarriot (2019). Also, see illustrations of this mixed evidence from Björkman and Svensson (2009); Arkedis *et al.* (2019); Mohanan *et al.* (2019); Raffler, Posner and Parkerson (2019) and Creighton et al. (2020).

[2] These intermediary non-governmental organized groups can take different forms, including civil society organizations (CSOs) and community-based organizations (CBOs), but their formal organization enables them to receive financial support for the implementation of a social accountability intervention. Other formal and informal groups might be included in or complement this category (faith-based groups, unions, consumer groups, professional associations, social movements, village committees led by non-state actors, etc.). International development organizations may partner with these or other types of groups. Recent research suggests engaging with target communities makes them more effective to implement pilot development interventions than other actors, but creates unique challenges to take innovations beyond the original site of innovation where they may lack similar conditions for quality implementation of interventions at scale (Usmani, Jeuland and Pattanayak 2021; also see Raffler, *et al.* 2019).

This contradiction links to a dilemma that puzzles practitioners: what can they do to take insights and lessons from promising localized (i.e. frontline) social accountability processes in the health sector to more people and places? While there is limited evidence available, research and practice over the last decade suggest that practitioners can pursue three common pathways to scale:

1) the replication of best practice, narrowly technocratic,
2) leveraging of countervailing power through resistance, and
3) based on resonance with state efforts, seeking joint gains despite mixed motives.

These pathways are different from each other because they view how change happens differently and they also conceptualize the central challenge to scale differently. One pathway sees change happen chiefly as a response to new information, another sees change happen chiefly through conflict, and another through deliberation, compromise, and collective action. As such, they place a different emphasis on the need for and dividends derived from conflict, and they take a different view of the potential of social learning.

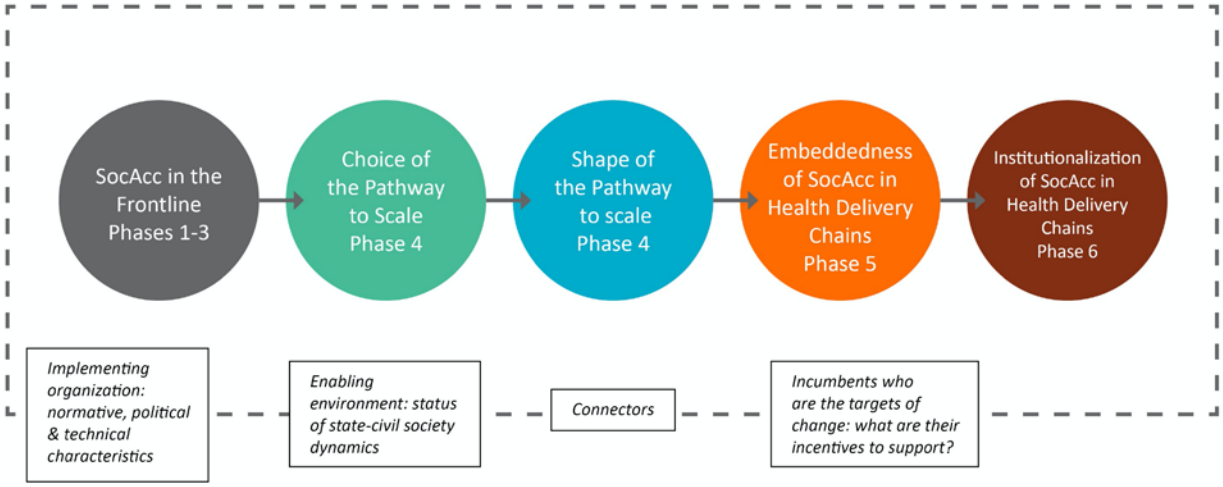
This study will combine these three pathways as part of the same research framework (a nested middle-level theory of change) and focus attention on the conditions under which each alternative pathway may be more promising to achieve greater scale. Since no single pathway is universally applicable; social accountability decision-makers need evidence to inform better decisions about who is well placed when and where to favor a path over another as well as under which conditions specific actors may benefit from changing course. This preempts a common problem within the social accountability field, when assumptions made within programs or evaluations may have only limited applicability outside specific areas of intervention and/or might mislead strategies and policy (Masset and White, 2019) [3]. To address this gap, the project will surface tacit knowledge from the social accountability field across a broad range of issues, including public sector dynamics and complex institutional change that demonstrate the range of types of paths. We focus on the health sector because there is a substantial body or research on social accountability processes in the health sector at the local level and there is also demand from decision-makers to understand the contradictions in relation to scale up.

Figure 1 provides the roadmap for this paper at a glance. It presents an overview of an indicative theory of change. This theory of change assumes that social accountability may contribute towards improved capacities to deliver health services and inform the scale up of those processes. In the health sector, most social accountability interventions are localized and start at the health service delivery (see Fox, 2015; ePact, 2016) (phases 1-3). Social accountability initiatives are complex interventions, often encompassing multiple components. Processes of service delivery can have different goals, take many forms, and engage various different actors. In theory, these processes often help surface problems that undermine effective delivery, either looking backwards to sanction those deemed responsible

[3] See box 2 for a definition.

for the problems of the past or with an eye towards prevention of those problems recurring in the future. Quick wins from problem-solving contribute to increased stakeholder alignment across the health delivery chain over time (phase 4).

Figure 1: Overview Theory of Change



When moving from the delivery of health services to the health sector policy arena at scale, there is no single pathway to embeddedness and institutionalization of social accountability (phase 4). For the purposes of this project, this is how we understand scale (Box 1 below) We map a number of options available to reformers to contribute to the *partial embeddedness* of social accountability in the health delivery chain (phase 5), and then to the *institutionalization* of social accountability in the health delivery chain (phase 6).

This theory of change nests two commonly used theories of change in the literature (best practice and resistance), along with the third pathway emerging from practice (resonance). We believe the inadequate theorization of this third pathway has created obstacles for knowledge accumulation and an evidence gap resulting from the differential speed at which monitoring and evaluation, on the one hand, and practice, on the other, have evolved. We argue that problem solving to improve service delivery, sector governance, and accountability or "collaborative social accountability (Guerzovich and Poli, 2020)" and resonance, which seeks joint gains despite mixed motives, are the key for effectiveness and scalability in a significant number of contexts.

Box 1: Two outcomes of scaling up social accountability: Embeddedness and institutionalization

This paper addresses two possible outcomes of scaling up of social accountability: **embeddedness**, in the short to medium term, and **institutionalization**, in the medium to long term. **Institutionalization** refers to those situations in which the principles and practices of social accountability are built in to the health system and structure interactions between actors over time. Institutionalization creates predictability and regularity for those interactions and the legitimate rules of the game (North, 1990). **Embeddedness** refers to the integration of social accountability principles into practice. For example, social accountability practitioners contribute to improving how health systems function by incorporating key principles of accountability within the health system so that the practice becomes an essential part of how those systems are designed in law and/or work in practice.

In many health systems around the world, social accountability is **not embedded**. For example, citizens can be invited by civil society organizations to participate in a social accountability process, receive civic education, participate in the collection of data on health standards, and engage in meetings with local health providers to share findings from monitoring. This typically happens in parallel with the workings of the health system and reform efforts implemented by other actors within the system, and the overall integration of principles of social accountability and practices is weak. For social accountability to be **fully embedded**: a) the state promotes, supports, or at least recognizes official spaces for citizen engagement in participatory processes that set communal priorities, plan health policies, design programs and budgets, execute (at least in part) policies and programs, often with the financial or technical backing of the state bureaucracy (management and delivery), and/or mechanisms to monitor and support the enforcement of sanctions, and b) these spaces are used in practice, contributing to learning, power sharing, and better health outcomes.

In between not embedded and fully embedded, there are multiple ways in which synergies can be established between social accountability processes and health delivery systems in limited ways: **partial embeddedness.** For example, SEND Ghana, a nonprofit organization, implemented a multistakeholder project, supported through a 4-year, $650,000 GPSA grant. Its objective was improving access and quality of health and education budget allocations, execution, and service delivery through iterative social accountability processes at central, regional, and district levels. Mills (2020) finds that SEND Ghana pursued a social accountability strategy to influence the health budget that originally worked in a scenario where demands were made by civil society in isolation, and the proposed citizens budgets were not deemed to be relevant by Parliament - demands were not integrated into policy. SEND Ghana then pivoted to a strategy led-by civil society in collaboration with the staff from the Ministry of Health. SEND has established formal agreements with relevant authorities at central, regional, and district levels and strengthened the capacity of SEND-Ghana's network of local civil society groups to coordinate and implement social accountability tools and processes in municipalities and health clinics. Multiple engagement actions were carried out and citizens' priorities considered in the budget and components of social accountability processes interact with health decision-making.

Embeddedness takes time and is a contextually contingent result, in the sense that interactions between actors can be ad hoc and easily reversed. Health programs that include social accountability principles can be adopted by one government but dismissed by the next. Laws enshrining formal entities such as village health committees may not be implemented in practice. Advances in embeddedness are often important, hard won battles but they are distinct from institutionalization.

This paper proceeds as follows. Section II introduces the study's key innovation: a nested middle-level theoretical framework to evaluate how social accountability might work and scale up in health service delivery.

Section III describes the three pathways we have identified from existing literature and practice in the social accountability field (resistance, resonance and best practice), and examines the assumptions underlining each path.

Section IV focuses on how social accountability processes in the health sector happen in practice in delivering services, identifying factors that evaluation and research evidence suggest matter for the effectiveness of those processes.

Section V zooms into the resonance pathway, which we argue is a promising pathway to achieve scale in various circumstances. Emerging from stakeholder engagement and practice, this pathway is a theoretical blind spot in the literature. For this reason, we feel it can help account for some of the apparent contradictions of mixed evidence and growing uptake. This section maps potential routes within this third pathway.

This focus on alternative causal pathways, along with the identification of the key actors and environmental characteristics associated with progress along each one, is key to providing practical cues for decision-makers to consider what resources are available to them, as well as the characteristics of the context in which they are operating. These considerations are critical for identifying which pathway may present the best opportunity to take social accountability from the frontline to scale in their own circumstances. The simultaneous focus on multiple pathways interacting with actors and environments with different characteristics can also help understand how much emphasis different pathways should be given in social accountability portfolios of interventions.

Section VI concludes by laying out the next steps for the project, including key methodological aspects of the approach to refine the middle-level theory proposed.

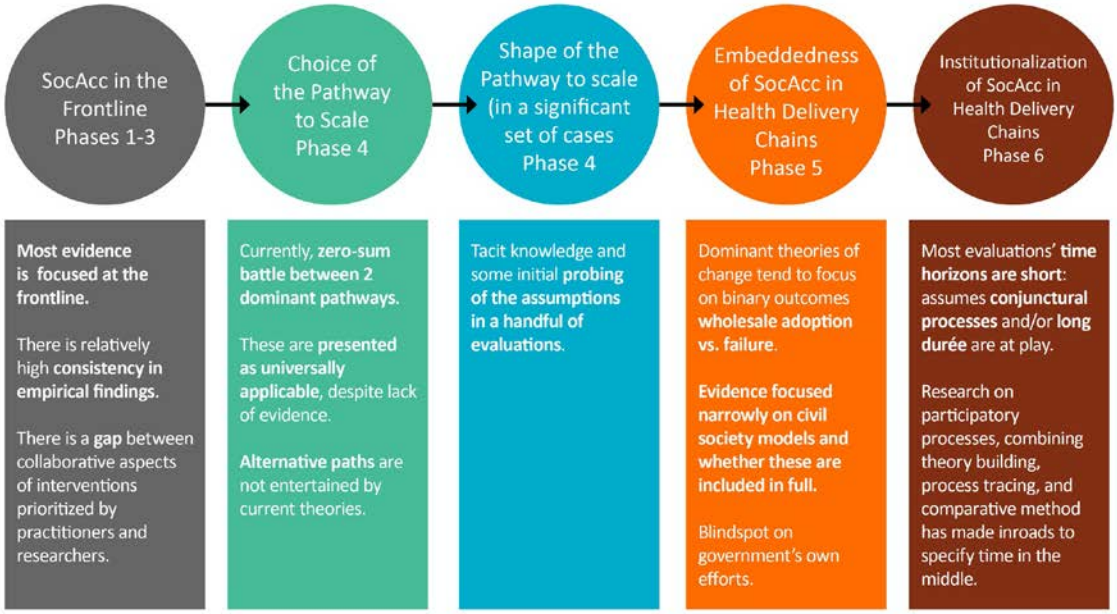
II. Innovations: A methodological approach to connect service-delivery to policy with theory

This section reviews knowledge from research, evaluation, and practice in the social accountability field, with a focus on the health sector (previewed in Figure 2). It examines knowledge gaps, including plausible reasons for these gaps and their consequences. Our analysis suggests that theoretical shortcomings are at the core of the growing chasm between knowledge and practice. We propose the development of a nested middle-level theory of change as a way to address this gap, building on Davey *et al.* (2018), Masset and White (2019), and Cartwright (2020)'s propositions regarding middle-level theory.

The choice to focus at the middle level takes into account a key characteristic of this body of evidence, which is not unique to social accountability. In the literature, the pathways

connecting social accountability results at the frontline with the policy arena are either: too abstract and general to help practitioners talk about, make informed choices, navigate and evaluate different options in concrete contexts (see the top of Figure 3); or too detailed and localized (see the bottom of Figure 3) to enable meaningful learning across social accountability practice (Sartori, 1970).

Figure 2: What do we know about social accountability?



The World Bank's 2004 World Development Report *Making Services Work for the Poor* (World Bank, 2003) changed the conversation regarding the delivery of public services, arguing for citizen-led transparency and accountability as a key feature of country and sector systems, including health. Since then, numerous evaluations have tried to test the effectiveness of social accountability interventions at the frontline, including many funded by the Foreign, Commonwealth and Development Office (FCDO).

While evidence of impact is mixed, systematic reviews show relatively consistent findings that many interventions achieve some form of positive results.[4] A substantial proportion of the literature has assumed that the mixed, localized results were partly the result of methodology, such as the quality of data available (see Joshi, 2014; Anderson, Fox and Gaventa, 2020), rather than looking for potential vulnerabilities in the assumptions which underpin mainstream theories of change. Theories in the social accountability field often fail to specify the key causal mechanisms linking citizen action to results in the health sector, or

[4] Reviews of evidence include Khemani, 2007; Rocha Menocal and Sharma, 2008; Gaventa and Barrett, 52010; McGee and Gaventa, 2011; Joshi, 2013; Westhorp *et al.* 2014; ePact, 2016; Fox, 2015; World Bank, 2018; Waddington *et al.* 2019; Tsai *et al.* 2019. Specific studies are discussed in the paper, including Björkman and Svensson, 2009; Raffler *et al.* 2019; Arkedis *et al.* 2019; and Mohanan *et al.* 2019.

where they are included, these are often assumed rather than empirically substantiated (Joshi, 2013; Grandovoinnet *et al*., 2015). Tsai *et al.* (2019) find that most of the interventions examined in the literature address only one or two steps in the causal pathways from generating new information to achieving governance outcomes.

Evidence about scaling up processes in social accountability and overlapping fields is also scarce and non-systematic (Adhikari et al., 2017: 12; also see Bold et al., 2013). Raffler, Posner and Parkerson (2019)'s replication at scale of Björkman and Svensson (2009)'s evaluation of the effects of mobilizing communities to use information in Uganda did not lead users to apply pressure on underperforming health providers. A randomized evaluation of another large-scale intervention in Tanzania and Indonesia had null results— meaning that, on average, the program did not measurably improve targeted health outcomes.[5]

Interpretations of these results are anchored in (and thus biased by) a recurrent debate between two theoretical-methodological positions. The first asks questions regarding measurable progress in human development and, consequently, and uses methods consistent with a focus on efficiency (see e.g. Tsai *et al*. 2019). The second asks somewhat different questions - emphasizing citizen empowerment and changing political relationships - and thus employs different methods to demonstrate this trajectory towards transformational change (see e.g. McGee and Gaventa, 2011). It is not surprising that these different starting questions shape very different narratives for what social accountability should be and what it ought to achieve.[6] While these two narratives about the field are not necessarily incompatible, each tends to present their propositions as if they were universally applicable.

This debate has led to a problem of conceptual stretching (Sartori, 1970). As we will discuss below, commonly used concepts in relation to social accountability scale up have been used to cover significant variation. Ideas of scale have been interpreted in various ways and apply to very different forms of social accountability.[7] This has come at the expense of clarity and utility, particularly in the case of the policy recommendations associated with those concepts. In these two camps, as USAID's David Jacobstein notes, "causal logics seem to have been implicitly transformed into paradigmatic and normative assumptions informing action and thus not sufficiently questioned by practitioners and evaluators in theories of change or tested empirically (Jacobstein's comment in Aston, 2020)." Amidst this lack of conceptual clarity and tendency towards theoretical confirmation bias, how would an evaluator

[5] E-Pact (2016), used a different methodology - Qualitative Comparative Analysis (QCA) – and also failed to confirm an alternative hypothesis about how scale happens. in a macro-evaluation of social accountability for DFID, was not able to confirm its hypothesis on vertical integration (feeding evidence and learning into higher level discussion and support for higher-level legislative change and policy change leads to improved higher-level [at-scale] service delivery). This suggests that a QCA may have been premature in order to answer questions in relation to scale. The use of theory-based methods such as that introduced in this paper may be required to determine necessary and sufficient conditions and how these may vary in different contexts.

[6] Briggs (2008) identifies the same dynamic in the debates among proponents of these logics in the debate about democracy.

[7] In addition to reviewing the literature, the team has reviewed and/or informed the operationalization of more than 2000 civil society programs in the last decade. We build on these insights as well.

confidently make recommendations beyond the specific programs assessed? How would a decision-maker know whether the findings were relevant to her circumstances?

Despite discussions about putting practitioners' theories front and center (Fox, 2019), the mental models that guide evaluation and research have created a self-reinforcing divide: a debate in which researchers and evaluators rarely ask about variations that are *not* at the core of the debate. Hence, practitioners rarely consider challenging the model by including other components of their work into evaluations (for an exception see Ball and Westhorp, 2018). This seems to inform which stories about social accountability get told, which data get collected, what research and evaluations are produced, which knowledge is included evaluation syntheses, and what questions garner repeated attention and inform investments and interventions in the field. This dynamic undermines the usefulness of theories to build narratives about "what the work is" as well as to guide monitoring, evaluation, research and learning (Jacobstein, 2020; Guerzovich, 2020).

The SALT project will redress this trend by focusing on competing assumptions in a single middle-level theoretical framework. This paper makes two contributions to this endeavor.

The first is to break down, prioritize and sequence the development and validation of multiple theories in the field of social accountability by nesting alternative sub- theories of change in an overarching theory. See box 2 on the next page for what we mean by sub-theories (or "nested") of change.

Existing evidence points to the importance of a focus at steps in the process between the front line and higher-level policymaking in a specific sector. As Levy and Walton (2013: 1) point out,

> *'These in-between spaces are major domains of political, stakeholder and organizational behavior. These are sources both of within-country and across-country variation in the quality of public service provision and also provide the locus where many opportunities for achieving gains in performance are to be found (Levy and Walton, 2013: 1; Levy, 2014).'*

The hope is that this will provide more understandable and practical guidance to support monitoring and evaluation of individual programs, while showing how different contributions across the field of social accountability fit together (see Mayne, 2015).

Box 2. Why a nested theory of change?

A "nested" theory of change is one that embeds different levels of detail (and abstraction) of theory (Pawson, 2013; Mayne, 2015; Punton *et al.* 2020) - from highly abstract universally applicable theories that are meant to organize a field of practice to intervention level theories of change passing by portfolio level (or grouping of interventions under the same umbrella).

Nested theories of change are especially well suited for sequencing and staging the complex set of interventions within the social accountability field. They allow us, first, to break down the field into simpler and more practical/manageable sub-theories of change that apply to groups of interventions. In this case, we identified 3 pathways to scale: the replication of best practice, leveraging the power of resistance, and based on resonance with state efforts that often inform portfolios of interventions. Each one of these pathways incorporates insights from highly abstract general archetypes which reflect general principles of human behavior such as response to information, carrots and sticks, or capacity building (Funnel and Rogers, 2011) or global theories such as "punctuated equilibrium," "policy windows," or "advocacy coalitions" (Stachowiak, 2013) commonly used (implicitly) in social accountability programing, research and evaluation. Instead of taking the universal ambitions of those sub-theories of change, and the associated zero-sum debate among them, at face value, a nested theory of change allows us to localize sub-theories in a framework that is well suited for asking when, how, and where the insights of those sub-theories may be a more promising bet than alternative pathways.

Second and related, a nested theory should allow us to investigate intervention-level theories to consider how well these may fit in these different portfolio-level sub-theories, including the enabling and disabling factors which help explain this fit. The goal of this sequencing and staging of interventions is to facilitate "an effective monitoring and evaluation plan that identifies what information is needed for each sub-Theory of Change and when (Mayne and Johnson, 2015: 414 cited in Richards, 2019)."

The second contribution is to identify for each one of these theories the causal principles that are applicable across different social accountability processes, recognizing that there are exceptions (Cartwright, 2020).[8] To achieve this, it is important to identify the priority contextual, procedural, and organizational factors that may support or undermine the effectiveness and sustainability of different pathways to scale. Following Cartwright, we call

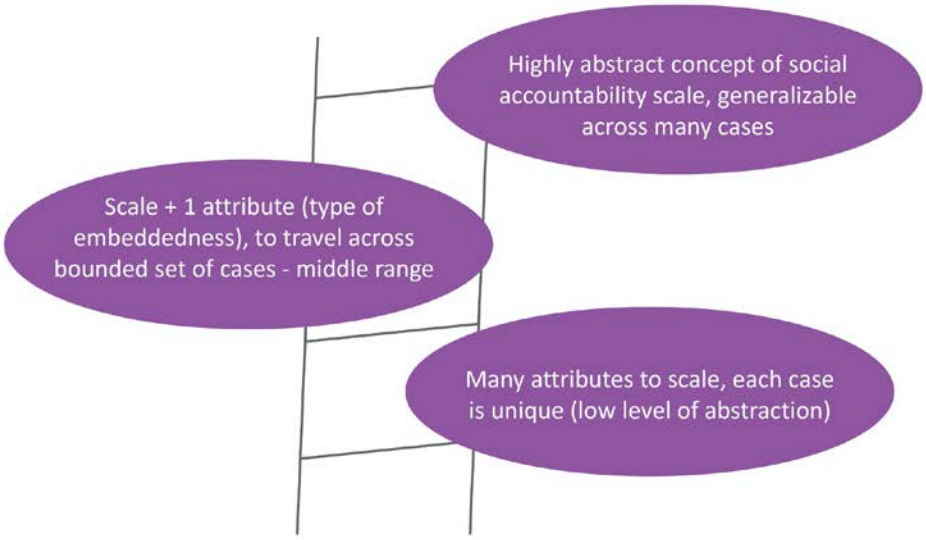
[8] The discussion so far helps to justify the choice for putting causal mechanisms between outcomes and interventions (Falleti and Lynch, 2009; Beach and Pedersen, 2019: 2 – 3, 31 – 32) at the center of research and evidence. Rather than focusing on 'nuts and bolts,' the body of knowledge should be focusing on the 'cogs and wheels' (Hernes, 1998: 78, in Beach, 2016: 465). To understand why some events unfold as they do, along with stakeholders' perspectives and interpretation of events, we need to focus on *how* things happen as much as what is happening (Falleti and Lynch, 2009; DFID, 2014: 7 - 8; see also Scriven, 2008: 20). In a field whose evidence base acknowledges that the 'how' (i.e. the process) matters to account for outcomes, yet no well-developed theory exists and there is limited prior knowledge, theory building in order to provide a basis for causal inference in theory-based methods happens through the identification of causal processes or 'chains' (Stern *et al.* 2012: 24).

these factors 'moderators.' Moderators are factors for which 'their absence or presence determines whether there is an effect or moderates the scale of that effect' (Cartwright, 2020: 5). Identifying moderators at critical junctures can be important in understanding whether the frontline action can produce the next step. The moderators we will examine include: 1) the normative, political and technical characteristics of the organization implementing the social accountability intervention, including its identity; 2) the incentives of incumbents who are the targets of change; 3) the presence and savviness of "connectors"; and 4) enabling environment (the status of state-civil society dynamics prior to the intervention).

We know that social accountability is a field in which context matters (Grandvoinnet *et al.* 2015; Tembo, 2012; Bukenya *et al.* 2012; O'Meally, 2013; Tsai *et al.* 2019); yet evidence from Randomized Control Trials (RCTs), which comes from very few countries (Tsai *et al.* 2019), along with evidence from qualitative research and evaluations (e.g. Joshi, 2017; Hernández *et al.* 2017; 2019; Anderson, Fox and Gaventa, 2020), has failed to produce conclusive evidence that may be transferable to other similar contexts. Given this, it is important to draw clear boundaries for the contexts and causal conditions under which each pathway may or may not be applicable.

An added benefit of this approach is that it provides an avenue to codify and embed insights from the emergence of "collaborative social accountability" (Guerzovich and Poli, 2020; Guerzovich, Poli and Fokkelman 2020; Poli, Guerzovich and Fokkelman, 2020) through the development of practice. Unlike the body of existing research (as discussed in Aston and Zimmer Santos, forthcoming), many social accountability practitioners appear to share the assumption that synergies between civil society and public sector reform efforts can be a win-win across the delivery chain, in terms of capacity for delivery (Ostrom, 1996; Doin *et al.* 2012; Levy, 2014; Guerzovich *et al.* 2020). Our approach embraces the concluding hypothesis in Waddington *et al*.'s (2019: 125) systematic review of the sector: 'citizen engagement interventions that do not incorporate complementary interventions along the service provider supply chain may be insufficient to improve key wellbeing outcomes for target communities.' Waddington and co-authors rely heavily on experimental methods, so process-related information accompanying these studies provided only partial insights regarding causal mechanisms. However, the hypothesis has been probed in several project evaluations combining process-tracing and comparative methods supported by the Global Partnership for Social Accountability – GPSA (e.g. Falisse *et al.* 2019; Mills, 2020) as well as e-Pact's (2016) review of DFID programing. The collection of findings provides important clues to: a) conceptualize outcomes of scaling up – focusing on adaptations, rather than binary adoption of wholesale civil society interventions or failure; b) focus on public sector actors' own efforts rather than only looking at those of civil society; and c) question the appropriate length of time, rather than assuming short-term time horizons or the long-durée (Nelson *et al.* 2018). These issues are discussed in greater detail in the remainder of the paper.

Figure 3: Moving in the ladder of abstraction



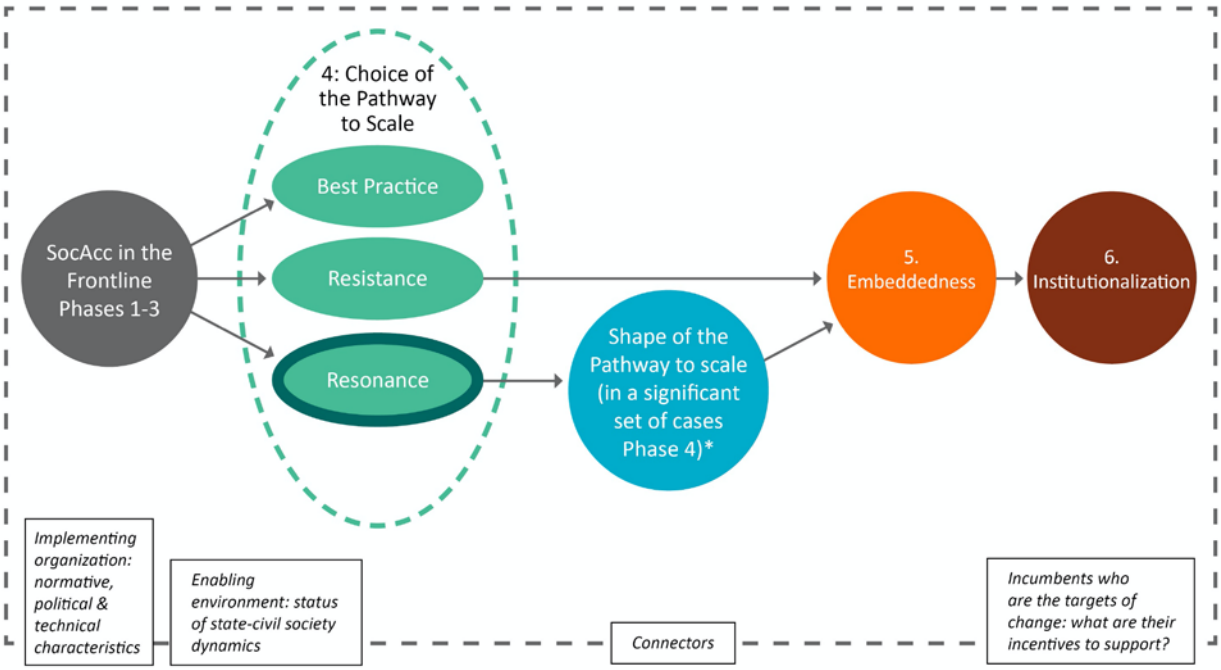
III. Choosing how to take lessons to scale

The next section considers how to choose between different pathways to take social accountability interventions to scale, beyond direct service delivery. Ideas about social accountability rarely, if ever, spread across the public sector without direct intervention. Specific individuals make a range of decisions to help identify, explain, and argue in favor of uptake. Other individuals will experiment, adopt, adapt and work around, negotiate and renegotiate, reject or ignore those insights. Others may block these processes.

Figure 5 illustrates a simplified sequence linking processes and results from the frontline (Phases 1-3) with results at scale (Phases 5-6). Phase 4, connecting these two spaces for action, synthesizes three related insights. First, there are multiple potential pathways to scale (resistance, best practice and resonance) – which is why a nested theory of change is needed. Second, scaling social accountability is as much about the pathways that accountability stakeholders pursue as it is about the outcomes. Third, key facilitating and blocking factors can help identify which pathway is ripe for action for a particular actor in a given circumstance. This informs whether actors should consider staying or changing course across pathways. The focus on multiple pathways interacting with different actors and environmental circumstances can also help understand how salient different pathways should be in social accountability intervention portfolios, rather than advocating a single

course of action. This theory of change is informed by these actors' agency and their perceptions, while also considering structural constraints.[9]

Figure 5: Indicative nested theory of change



We focus on two challenges to be addressed in going to scale that are central in the literature about institutional change. These are: a) the perceived resistance to change by the "other" (inside or outside the state) and the status quo interests that preempt change and, b) the challenge of social learning understood as including stakeholders who learn to co-operate to resolve collective action challenges in a "win-win" way. The resistance, best practice, and resonance pathways emphasize different combinations of these challenges. The discussion underscores the potential of the resonance pathway, often overlooked in the literature but not in tacit knowledge.

The resonance pathway

For the resonance pathway, the central challenge is "social learning" in the sense of enabling a group of individuals to organize and work out how to make the most of a situation to create shared gains, even when there are conflicting interests and all face

[9] As discussed below, these insights are also consistent with theories about the diffusion and scale up of complex interventions in the health sector in high income countries, including Greenghlagh *et al.* (2004 and 2017). This theory of change builds on those insights, while remaining anchored in theory and practice in the social accountability space. The authors thank the paper's anonymous peer reviewer for pointing us to common threads across both bodies of work.

temptations to free ride, shirk or act opportunistically (for a canonical statement on the evolution of institutions for collective action, see Ostrom, 1990).

Resonance focuses on addressing collective action problems in a predominantly game theoretical sense (Ostrom, 1990; also see World Bank, 2017). Stakeholders who adopt this route are aware of power asymmetries and the role of institutions as distributional instruments loaded with power implications and of the potential 'gap between normative expectations and empirical realities (Hall, 1986; Thelen, 1999; Mahoney and Thelen, 2010; Booth, 2012; Cornwall and Coelho, 2007; 5).' They also tend to consider that bureaucracies include rent-seekers, but also mission driven bureaucrats (Honig, 2018). From this perspective, motivating and empowering bureaucrats and relational forms of accountability may create a more promising avenue to develop joint capacities for action, improved governance and service delivery than only focusing on oversight and sanctions vis-a-vis agents in the public sector (Tendler, 1997; also see, Briggs, 2008; Honig and Pritchett, 2019).[10]

This understanding of collective action, therefore, differs from a social movement understanding of collective action found in a resistance pathway presented below, which focuses on networks of empowered citizens tapping into critical junctures to channel grievances or desires to state actors (Green, 2017: 14), and with support from 'external allies' may gain 'sufficient clout to influence public sector performance' (Fox, 2015: 352). Both pathways recognize the importance of power and politics yet have different normative expectations and make different calculations regarding the potential of zero-sum adversarial politics versus the potential for reaching shared understandings and opportunities for obtaining gains from cooperation.[11]

The linchpin of the resonance pathway is eliciting a scale up by increasing the alignment of targeted stakeholders so that they identify and pursue a functional fit for how social accountability helps policymakers solve concrete problems within the chains of health provision (Levy, 2014; Guerzovich and Poli, 2020). This approach puts similar emphasis on conceptualizations of social accountability processes as "deliberative spaces" where demands are registered, debated, mediated and justifications are produced, rather than citizens demanding a response from providers (Teskey, 2021).[12] However, multi-stakeholder commitment, coordination and co-operation to address governance failures for improved health delivery are rarely easy (World Bank, 2017).

[10] These arguments are consistent with insights in a literature about the public administration literature on bureaucratic motivation that is too broad to cite here as well as with works about different forms of accountability such as Behn (2001).

[11] These two pathways also embody, often implicitly, two alternative positions about how institutional change happens - a body of work that is too extensive to discuss here. The resistance pathway shares many characteristics of institutional change explanations that combine critical juncture and path dependence frameworks. Hence, it can be "both too contingent and too deterministic," (Thelen, 1999: 385). The resonance pathway conceptualizes presumed institutional stability as grounded on ongoing active adaptations to the environment and institutional change as gradual and piecemeal, even during crucial junctures (Thelen, 1999; 2004; Mahoney and Thelen, 2010).

[12] Teskey (2021) notes the emergence of this conceptualization of social accountability, consistent with literature focused on the "middle" of the state-society relationships and collaborative social accountability frameworks cited in this document. It is important to note, however, that this re-emergence and link to the literature on deliberative processes and participatory democracy was part of early theoretical debates.

Actors inside and outside the public sector have to share an understanding that there are options with higher returns for all (Fukuyama 2004; Ostrom, 1990; Ostrom, 1998). They also need capacities to devise workable agreements and implement those agreements together (Briggs, 2008). It is important to underscore that obtaining shared gains (i.e. scalability) does not mean that these actors reached a decision about the best possible course of action by consensus, their interests have converged, or asymmetries of power and conflict have disappeared. Scaling through collective action requires both social learning and bargaining among actors with different interests and power.

While a tradition of civic cooperation, stocks of shared norms such as reliability and reciprocity, capacities to work with others, networks, and trust (social capital) can foster collective action (Putnam *et al.* 1993), in many contexts where social accountability has the potential to scale those assets and the capacities associated with them are scarce (Guerzovich *et al.* 2020; Poli *et al.* 2020).

The literature on collective action shows that interactions among individuals can contribute to sharing information, learning new norms, developing trust and capacities to get things done by working with other actors, and delivering and sustaining win-win results (Poteete, Janssen, and Ostrom, 2010). The literature on scale up of innovations in service delivery have found that external change agents are more likely to contribute to that process if they explore and emphasize with potential adopters' perspectives, among other characteristics associated with the development of strong interpersonal relationships (Greenghlagh *et al.* 2004). In particular, interactions through processes of collaborative social accountability (Poli and Guerzovich, 2020) and co-produced accountability (Doin *et al.* 2012; Guerzovich and Schommer, 2016) can contribute to shaping different incentives, beliefs, preferences, forms of contestability and relationships of accountability among stakeholders. Certain forms of external support might be able to create more conducive conditions and/or enable the development of capacities for these collective results (World Bank, 2017; Guerzovich *et al.* 2020; Guerzovich and Poli, 2020).

Practitioners that pursue this pathway are thinking and acting politically to work around a situation that does not lend itself to scaling social accountability into one that does via collective action. They need to figure out collectively how to do so feasibly and effectively (setting the direction and implementing it). Generally, with a deep knowledge of barriers built into the status quo, they assess that confronting potential opponents is not necessarily the shrewdest course of action. The more productive use of their capacity to generate results might be to subvert the status quo by taking advantage of the openings, dynamic tensions and pressures for change built into the health system and its institutions (Mahoney and Thelen, 2010; Levy and Walton, 2013).

There are many plausible routes through which different actors can work around the status quo and make (and learn to make) social accountability part of the suite of interventions that are part of the health sector's work, when they perceive that social accountability contributes to solving problems pertinent to them. The next section discusses these pathways in detail. To the naked eye, much of this action, deliberation and compromises can be perceived as supporting institutions, but in "working with the grain" of the system changemakers hope for

and often nurture gradual change from the periphery without the need to break the rules (Mahoney and Thelen, 2010; Levy, 2014).

The resistance pathway

The conceptualization and operationalization of this pathway often starts with an assumption of an adversarial relationship of resistance between 'us' and 'them': civil society actors and status quo interests entrenched in the health system and the state (see Fox, 2016). Associated to this is the belief that change happens through productive conflict (resistance, tension, pressure) between both sides, and typically that the primary motor for scale up is civil society agenda setting in spaces created by coalitions of relatively powerless or excluded groups.

Theories in this category assert the goal of squeezing anti-accountability forces by putting them under pressure to commit to change. It is this pressure which is deemed to give civil society demands sufficient clout to deter poor behavior and incentivize stakeholder compliance locally and at scale (Joshi, 2013; Fox, 2015; Anderson, Fox and Gaventa, 2020). Adversarial "countervailing power" by civil society through multi-pronged campaigns commonly includes some form of "naming and shaming" from campaigners and potentially also including litigation and protest at different administrative levels is seen as key to success (Gaventa and McGee, 2010; Joshi, 2013; Fox, 2016; Joshi, 2017). Social learning is shaped by othering: relationships of resistance are mutually constituted (see Feruglio, 2017). As they do so, stakeholders learn about the struggle, whether this is within the state or within civil society.

Civil society actors may also have public sector allies on the inside (pro-accountability champions), including Supreme Audit Institutions (SAIs), to give "teeth (Fox, 2015)" to civil society resistance (see Halloran, 2015; Fox, 2016), yet these actors are viewed as a potentially ostracized minority of entryists and dissidents contravening established government norms and rules (Gaventa and McGee, 2011; Fox, 2014).[13] However, a growing body of evidence suggests that it is difficult to redeploy adversarial forms of countervailing power into collaborative ones which may be required to make reforms sustainable (Fung and Wright, 2003; Larsen, 2015; Feruglio, 2017; Edwards *et al.* 2020; Poli *et al.* 2020; Aston and Zimmer

[13] The literature anchored in the resistance pathway is more populated than alternatives in the one in the social accountability space. Authors in this tradition have discussed, although not always fully specified, alternative sub-pathways and the conditions under which each of these sub-pathways may be a plausible and/or better bet. For example, Fox (2016: 22) argues that it is possible to have more and less productive 'deployment of adversarial processes' through what he terms as a 'three-dimensional' understanding of constructive and confrontational forms of engagement (Fox, 2016: 22). However, if 'targeted conflict [is deemed] necessary to produce accountability (Fox, 2016: 23)' and if efforts are commonly interpreted by state counterparts as being confrontational anyway (Fox, 2016; Feruglio, 2017), then the net effects of more confrontational and "hybrid" approaches may not be very different in practice. Other literature in this and other pathways' tradition provides alternative arguments to justify the usefulness of the analytical dichotomy between resistance and cooperative approaches to explicitly explore alternative configurations and the contexts under which they may operate instead of Fox's preferred definition of approaches as part of a continuum (see e.g. Fung and Wright 2003; Kosack and Fung 2013a; 2013b; 2014). At this stage in the theory building exercise, the paper identifies this pathway and allows for gradients of it (see Table 1) to co-exist, and potentially interact with specifications of other pathways under concrete conditions that may affect their effectiveness.

Santos, forthcoming).[14] This raises questions regarding how generalizable this pathway may be across different contexts and in what concrete circumstances the pathway has greater potential to contribute to scale than alternatives identified.

The best practice pathway

The best practice pathway, conversely, is less attentive to either the challenge of perceived resistance by status quo forces and social learning. For years, many stakeholders' approach to social accountability scale up, much like in traditional development approaches, largely focused on designing best-practice solutions. This approach to scale up assumes that technical experts who produce knowledge can determine the unique form of social accountability mechanisms that work and then use their authority and knowledge to promote cross-context convergence (i.e. scale up) towards those arrangements (Rodrik 2008). Forms of "rigorous" knowledge used for this purpose can include transnational measurement and benchmarking and comparisons in different settings to incentivize a race to the top and randomized control trials, among other tools.

While this "first-best" pathway to scale up interventions has been frequently assumed in social accountability fundraising strategies (see e.g. Guerzovich and Poli 2014), it is less frequently made explicit in the social accountability literature. We can draw on research on Development Innovation Ventures (Duflo and Kremer, 2015; Kremer *et al.* 2019) to illustrate the logic that privileges best practice. That research assessed innovations that scaled, including those in social accountability in the health sector such as community monitoring to improve health delivery in Sierra Leone and digital attendance monitoring in India.[15] Kremer *et al.* (2019) found that although innovations often took a decade or more to scale, they found a strong correlation between rigorous evaluation and scaling (typically RCTs), argued that it was important to include academics in the design of models to improve innovations over time, and 'those that included development economics researchers were six times more likely to scale than those that did not (Kremer *et al.* 2019: 3),' and typically rely on building local technical capacity to bring evidence to bear on policies.[16] Neither politics nor power nor other conditions for implementation were considered of significant importance to whether innovations scale criticized in the broader governance and development literature (Levy

[14] In fact, references to "coalitions" and "partnerships" in this body of work, amidst accounts of adversarial stand-offs, rarely tell us much about the compromises needed for working with others (Briggs, 2008). Recent theory building efforts for hybrid approaches suggests that grievance redress is the weakest part of social accountability efforts (Joshi, 2017) and global evidence indicates that citizen engagement is the weakest component of social accountability efforts (Mendiburu, 2020; Mills, 2020).

[15] There are other sub-pathways in this tradition, including, in the team's experience, many embedded in implicit and explicit theory of actions to justify funding, monitor and evaluate research of social accountability at scale via randomized controlled trials through projects such as Raffler *et al.* (2019) or Arkedis *et al*. (2019) as well as other interventions that promote "races to the top" via international data and benchmarking, often complemented with technical assistance and/or other resources (funding and technical credibility , etc.) from experts. Much like in the case of the resistance pathway, the theory of change presented here allows for sub-pathways, including hybrid ones (see Table 1), without fully mapping all possible specifications at this stage in the theory building process.

[16] Kremer *et al.* (2019: 42) note that digital attendance monitoring 'did not scale beyond the initial RCT [but] the RCT itself was conducted at scale.' The authors also found a strong correlation with low costs, but not for the social accountability interventions included.

2014; Andrews, Prittchet and Woolcok 2017, World Development Report 2017, among others).[17]

Some stakeholders focus on piloting social accountability in spaces that are somewhat autonomous from the health system. Think, for example, of civil society actors coming together, designing a scorecard process, and monitoring the performance of doctors and nurses. If they follow up this intervention, it is by engaging those doctors and nurses after they analyze their data, also called 'transparency plus interventions (see Tsai *et al.* 2019).' If and when these local experiences produce results and those results are validated through research and evaluation, best practices are identified. Social accountability constituencies then deploy communication strategies to inform health managers with the power to augment best practices into larger scale interventions, for example by following the same process in new geographical sites (see Duflo *et al.* 2012; Gaduh *et al.* 2020). The simplification of complex interventions, that are broken down into more manageable parts (i.e. focusing on the transparency and/or the data rather than the plus) can create the perception that they are easily scalable (see Greenhalgh, *et al.* 2004). We see this in the way that participatory budgeting in places such as Mexico or Peru has been imposed from above, often as the result of the adoption of narrowly defined technocratic best practices. We also see it in cross-country benchmarking of citizen driven data about service delivery governance and outcomes, including the sustainable development goals. In Uttar Pradesh, for example, the World Bank is supporting a process to test and replicate social accountability interventions with support from the government, with no organized demand or leading role from civil society. This is best suited to contexts in which an evidence base is clear, and solutions are apparently well known, and thus where most of what is required is "packaging" evidence to suit audiences (see Oliver, 2019).

The key challenges for each pathway are displayed in table 1 on the next page:

[17] Much like the resonance pathways, there are variations across the pathway and evolutions over time which move away from the "pure" type, where academic institutions embrace the role of knowledge mediators and other political tasks associated with uptake and sustainability (see e.g. Gugerty and Karlan 2018; McAnnaly-Linz et.al. 2021). In many cases, the "expert" power of academics is complemented by the advisory roles, capacity building financing and other forms of power exercised by development partners who can also legitimize support and/or incentivize the uptake of knowledge via advisory services and other instruments (see e.g. Knack *et al.* 2020) and/or via teaching and judging international norms through technocratic instruments (for a critical outlook, see Broome *et al.* 2017). This is a moderator for these and other pathways that the project will explore below.

Table 1: Key challenges for each pathway

<table>
<tr><td rowspan="3">Perceived opposition</td><td>High</td><td colspan="2">Resistance</td><td></td></tr>
<tr><td>Medium</td><td></td><td></td><td rowspan="2">Resonance</td></tr>
<tr><td>Low</td><td>Best practice</td><td></td></tr>
<tr><td></td><td></td><td>Low</td><td>Medium</td><td>High</td></tr>
<tr><td></td><td></td><td colspan="3">Role of social learning</td></tr>
</table>

It should be underscored that the project's goal is to develop a nested middle-level theory of change, as discussed above. This has three important implications for the interpretation and use of these alternative pathways. First, while proponents and practitioners working within each pathway have differences, the focus here is at a higher level of abstraction to focus attention on the conditions under which each alternative pathway may be a more promising bet. The trade-off is less attention to these internal divergences and their applications.

Second, the rationale for building a nested theory of change at middle-level is that no single pathway is universally applicable. The interest is in better understanding in which specific contextual configuration each pathway may be a better bet. In betting on one pathway or another, politically savvy, evidence-informed actors will consider what resources are available to their organization, as well as which pathway has greater potential given the characteristics of the system in which they are operating (that is, the moderators at work). Different actors working at the same time in the same system may make different calculations and reach different conclusions about the most promising pathway to scale for their interventions. An actor or collection of actors may prioritize one pathway for scaling an intervention at a point in time, but another one as contextual conditions change.

Third, from an evaluation and research perspective, analysis of the theory of change of a collection of interventions calls for systematically questioning and comparing whether and how these pathways are combined across space and time. A theory of change, therefore, has to include and specify the relevant alternatives and the conditions under which they are expected to operate to enable systematic inquiry that can inform decision-making at critical junctures in the causal chain.

Table 2 introduces the key actors, actors' characteristics, and enabling environment that might explain whether changes take place along the resistance, best practice and resonance pathways and whether they are reverted over time. We prioritize these moderators because we suspect they can help understand when and where a particular pathway is more promising in two critical junctures in the causal chain where changemakers can make strategic decisions about their course of action. While the same moderators matter for all pathways, their values and direction can vary across pathways.

Table 2: Actors and environment: **key attributes for success, by pathway**

	Resonance	**Resistance**	**Best Practice**
Implementing organization: normative, political and technical characteristics	Readiness to collaborate with state actors, work "with the grain," and have an understanding of the inside of the sector	Readiness to openly confront state actors, e.g. to "name and shame" and have (long) established competence on campaigning repertoires	Readiness to produce data and evidence and work with counterparts to disseminate more efficient innovations over time
Connectors	Legitimacy to convene stakeholders in different worlds into a joint conversation with a stock of institutional memory of lessons from the past	Legitimacy to understand and represent grassroots and elites; *and/or* Legitimate actor(s) in the public sector who cut ranks, at potential career risk	Researchers and others with technical and academic credentials with ties to the region
Incumbents who are the targets of change: what are their incentives to support?	Short- and medium term political and reputational dividends	Window of opportunity - short-term political incentives, even in unlikely contexts	High rate of return on investments for government adopters and high social benefits for successful innovations
Enabling environment: status of state-civil society dynamics	Prior experience of deliberation with other types, especially in the sector's ecosystem	Created spaces and/or inclusiveness of invited spaces, especially in the oversight ecosystem	Monitoring and evaluation infrastructure from government, academia and/or civil society.

The interactions of causal mechanisms (Table 1) and context (Table 2) informs this project's working hypotheses:

1. Implementing organizations/actors have cultures which predispose (or bias) their interventions towards best practice, resistance, or resonance - "if you have a hammer, you see nails everywhere."
2. Each approach will succeed if the context (binding constraint) is indeed the one for which its preferred approach is well-suited to address (and vice versa).
3. A further necessary condition for success (given that the fit with context is right in the first place) is that the actors have and/or develop the right capacities and skills (hammer) for the selected strategy.

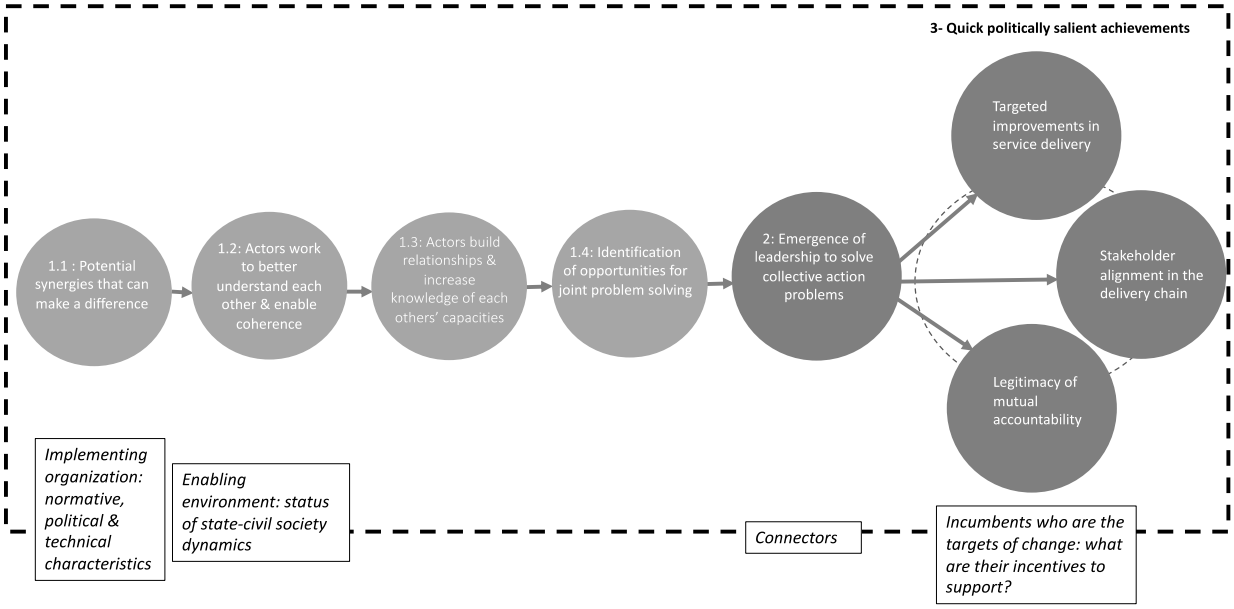
IV. Learning by doing at the frontline

This section focuses on insights across pathways of social accountability in the frontline, with special emphasis on the resonance. The importance of considering a pathway for scale up from collaborative social accountability at the frontline to other sites through resonance is consistent with a growing body of evidence, as well as tacit knowledge from practice. On the one hand, efforts to threaten and impose social or formal sanctions come with potentially significant risks for a wide variety of stakeholders (see Aston and Zimmer Santos, forthcoming). When given a choice, communities often opt for collaboration over confrontation (Kosak *et al.* 2019, Cant and Sarriot, 2019). Various studies of social accountability in the health sector also reveal that service providers' and managers' appetite to enforce sanctions may be quite limited (Banerjee *et al.* 2007). Nxumalo *et al.* (2018: 5) have argued that the 'laudable intentions' of imposing formal sanctions may have 'unintended consequences,' distracting staff attention, diminishing staff motivation and even weakening trust between providers and the community. And indeed, few of these interventions have demonstrated sustainability (Aston and Zimmer Santos, forthcoming).

In our review of existing literature, we found that *empirical* evidence for the relatively greater effectiveness of sanctions in influencing outcomes is either absent of thinly concentrated in few contexts and in programming which often *lacks* a citizen engagement component (Rocha Menocal, 2008; Gaventa and Barrett, 2010; McGee and Gaventa, 2011; Fox, 2014; e-Pact, 2016; Tsai *et al.* 2019; Waddington *et al.* 2019). On what appears to be normative rather than empirical grounds, the authors lament that studies emphasized the 'need to collaborate and support service providers and local officials in place of a more adversarial approach that threatens sanctions and reputational risk (Boydell and Keesbury, 2014: 18 also see Tsai *et al.* 2019).' Evidence from South Africa, Kenya, Peru, and Nepal's health sectors all suggests that such a collaborative approach with public sector actors is an enabling feature for delivering outcomes (Aston, 2015; Nxumalo *et al.* 2018; Gurung *et al.* 2019). Collaboration matters because, often, individual citizens and service providers cannot deliver all results on their own. Yet, civil society groups make decisions as to whether or how to incorporate this insight into their interventions.

Figure 4 presents an indicative theory of change at the frontline and corresponds to section 1 of Figure 2 (phases 1 – 3). It maps how social accountability processes might facilitate multi-stakeholder collaboration, resulting in leadership to address problems in the delivery of health and, improve results. The next section explores the steps and assumptions in this theory of change, drawing on examples and evidence from the implementation of social accountability interventions in the delivery of health services.

Figure 4: Theory of Change for Social Accountability at the frontline



1.1 *Finding potential synergies that can make a difference*

Civil society actors who embark in social accountability processes have a choice – which can be implicit or explicit. Their intervention can be designed as a standalone process that only reflects their preferences, regardless of what other actors are doing. The option is to identify what others are doing, including what reforms public sector actors are implementing, and design an intervention that tries to build synergies across efforts so that the whole is more than the sum of the parts.

1.2 *Public sector and civil society actors work to better understand each other and enable coherence*

When CSOs opt to build synergies with providers and public sector reform efforts early on, citizens and CSOs may be in a position to create a three-way process in which the problems that are salient for state actors become part of the conversation by design. Non-governmental organized groups progressively improve their capacity to identify citizens' problems and

citizens increase their receptiveness to the ideas brought in by other groups[18]. Adding the voices of those in the public sector has an added advantage: it is easier to identify what protocols, procedures and directives can provide the authorization space to address the identified problems and, in turn, how civil society demands might be more effectively anchored in the public sector apparatus. The introduction of so-called "interface spaces and processes," through which different stakeholders engage, can also strengthen the possibilities to identify problems and ways forward collectively.

1.3 *Public sector and civil society actors build relationships and increase knowledge of each other's capacities*

CSOs seek to understand the interests and incentives of frontline service providers and the opportunities and constraints they face. With greater receptiveness and a goal to build a common understanding, service providers communicate more openly regarding the resources available for delivery and capacity constraints and commit to resolve problems within their control.[19] Regular engagement means that stakeholders may gain a clearer understanding of what their interests might be, what motivates each other to collaborate, as well as learn to work together. However, features associated with the design and goals of the social accountability process affect the extent and depth of relationship building.[20]

1.4 *Identification of opportunities for joint problem-solving*

With a greater sense of mutual understanding, respect and joint capacities, service providers and organized civil society identify ways to address resolvable problems together and what mechanisms are required to resolve problems beyond their control higher up the delivery chain. Here, an intervention led by a non-governmental organized group might build citizens' and providers' capacities to access and/or reallocate resources available to them, while

[18] Empirical research in the field suggests that this is important because when citizens only learn what providers and officials *should* deliver according to rights, policies and standards, rather than also learning what providers and officials *can* deliver, given constraints and challenges, social accountability processes may create conflict, damage relationships, and result in repercussions on service users themselves (Banerjee *et al.* 2007; Aston, 2015; Lodenstein *et al.* 2018; Gurung *et al.* 2019; Squires *et al.* 2020; Hernández *et al.* 2020). Nxumalo *et al*. (2018: 12) further illustrate these multi-actor dynamics in Kenya and South Africa's health sector, which showed that 'managers gravitate towards more relational processes of accountability' and 'collaborative goal-setting to support shared outcomes' to reach a 'shared understanding.'

[19] Evidence from Peru, Nepal, and Mozambique shows this is an important step (Aston, 2015; Gurung et al. 2019; Schaaf et al. 2020).

[20] For example, in Indonesia, Ball and Westhorp (2018) found that building a social accountability annual cycle can pay off in the effectiveness of the processes and service delivery in later years. They explain: "experiences of receiving constructive feedback and improving services in response to feedback increases leaders' understanding of the value of citizen input, which encourages leaders to sustain citizen-based monitoring over time." In the Democratic Republic of the Congo (DRC), Falisse *et al.* (2019) found that a similar approach triggered a virtuous cycle of collective action in Village Health Committees in one province, where organizations are in a "learning phase", trying to learn from each other and collaborate rather than compete. However, this effect was not seen in other provinces. Plausible causes include the long-term efforts by external facilitators, who can leverage relationships built across project cycles to facilitate coordination and joint action, in the former but not the latter provinces.

improving the mutual understanding of those problems that cannot be solved within the parameters of the health delivery system (Vargas, 2013; Aston, 2015; Walker, 2018).

Phases 1.1 – 1.4 can be valuable in and of themselves, as documented in evidence syntheses of social accountability's localized results (e.g. e-Pact, 2016; Waddington *et al.* 2019). At the local level, alternative components of the social accountability process can produce a range of results from individual empowerment, to new information, to greater awareness about health standards, among other results that the interaction between CSOs and citizens through social accountability processes can create.

2. *Emergence of leadership to solve collective action problems*

This step in the theory of change assumes that **leadership emerges that can help address problems identified through the social accountability process.** This is a key intermediate step towards delivering those results on the frontline, eventually triggering changes in behavior from different actors.[21] Who are these leaders? They are often directly involved in service delivery. For example, health facility managers can stretch and adapt protocols, procedures and directives to solve low hanging fruit problems broadly within their control (e.g. targeting of capacities, responsibilities, assignment of tasks, and mandates and/or creating new capacities, redirection of resources, division of labor, and reinterpreting mandates in an expansive way). As they implement the intervention, leaders will often reinvent and adapt the intervention and put existing institutions to work.

3. *Quick politically salient achievements*

What results can this leadership unlock? Phase 3 refers to **quick but politically salient achievements** in health care (i.e. viewed as relevant and of some benefit to decision-makers). First, there are concrete improvements in service delivery, including improving standards of care, for instance. Second, leadership emerging from joint action of different stakeholders can increase the legitimacy of social accountability processes for solving problems and can also generate related results such as increased trust among civil society organizations, citizens, frontline providers and some managerial level actors, who also act as connectors to other stakeholders within the service delivery chain. Third, this kind of leadership can facilitate the

[21] For example, Long and Panday (2020) found that a social accountability intervention in Bangladesh contributed to service delivery results by creating spaces for multi-stakeholder dialogue, sparking efforts to overcome collective action problems at the local level and encouraging actors to overcome inertia - all without altering public officials' formal scope of work or mandate (also, see comparable examples in Indonesia and DRC: Ball and Westhorp 2018; Falisse *et al.* 2019). Collective action, which does not depend on the number of actors mobilized but rather on their stake, standing, and relevant resources (Levy, 2014; Guerzovich and Poli, 2014), is what may help turn localized results into inputs for other processes beyond the initial health centers and villages. In the Bangladesh case, this included filtering up the system insights from the frontline so that actors with mandates to help solve problems at the subnational and national level might contribute to targeted problem solving by unlocking existing mandates (Long and Panday, 2020). The qualitative analysis of scorecard programs in Tanzania, Indonesia, Malawi, Sierra, Leone and Ghana found that involving officials who have responsibility and authority over district services as brokers of state-society collaboration around a shared goal of improving a valued public service offers opportunities for mutually constructive collaboration, making social accountability interventions more effective (Kosack et *al.* 2021).

alignment of those serving on the frontline and their managers, cultivating alternatives as a foundation for action and mitigating the challenge of administrative capacities that are too thinly distributed across the public sector (McDonnell, 2020).[22]

As frontline public sector actors and civil society representatives increasingly see the benefits of collective action, they can demonstrate their own resourcefulness and performance to their managers. They also reach out to other actors in the delivery chain, especially to support frontline problem solving. In turn, actors at managerial and governance levels may develop a favorable attitude towards collaborative social accountability.

Mid-level managers can take advantage of moments of relative institutional, policy or programmatic indeterminism, by bringing their own concrete proposals (Levy and Walton, 2013)[23]. As social accountability informs officials of local realities, they become better able to mobilize resources to improve delivery, respond to human resource problems in the public sector, and seek improved assistance from line ministries.[24] In fact, means of re-alignment of actors in the system might, often subtly, reflect renegotiations of power relations between the State, civil society organizations, and citizens.

Conversely, where the above discussed results (as pictured in Figure 4) do not materialize, the building blocks that combine to establish a firm foundation for scale are absent. We propose that this may, in part, account for the lack of results in many social accountability interventions and the apparent failure of to scale up that has been documented in the literature (e.g. the lack of scaling up in the health sector is addressed by Greenhalgh *et al*. 2017).

Lastly, we assume that how action is designed and implemented at the frontline (Phases 1 - 3) matters to understand which actors learn about the value of social accountability and how social accountability's learning, processes, and/or value addition may be scaled up and sustained over time. Whether the social accountability process has been chiefly focused on producing citizen-generated data to inform decisions, facilitating dialogue and problem-solving between different stakeholders in the health system, and/or redressing grievances, can have important implications for how actors understand and advocate for social accountability in other sites in the delivery chain. What did stakeholders learn about social accountability's main function? (This could include to facilitate transparency, collective action for problem solving, and/or sanctions). How do leaders effectively convince decision-makers that may take up social accountability lessons beyond the original site of action? The next section looks into these questions.

[22] Recent evaluations of collaborative social accountability processes in health in Indonesia, DRC and Bangladesh underscore the role of these processes in overcoming inertia, catalyzing multi-stakeholder collective action and, in so doing, strengthening the health system (Ball and Westhorp, 2018; Falisse et al., 2019, Long and Panday, 2020).

[23] This engagement of the managerial and governance levels can help overcome "low accountability traps" (Fox, 2014) because these actors bring additional capacities, mandates and resources to the table.

[24] For example, social accountability processes can contribute to results partly by 'directly and indirectly activating and motivating other actors, institutions, and structures (like pre-existing committees and groups) and then bolstering them and their efforts' (Long and Panday, 2020: 56).

V. What does the resonance pathway look like?

This section maps different forms the resonance pathway can take in practice, including what embeddedness may look like (see Figure 6). In these processes, which often occur outside the political spotlight, ownership is transferred from CSOs to government insiders (Gbeleou and Schechter, 2020). Solutions result from complex chains of decisions, and from intended and unintended consequences that are the product of the interaction of social accountability proponents, coalitions blocking social accountability, and a broad range of factors beyond the control of those actors, in particular institutional contexts, over time.

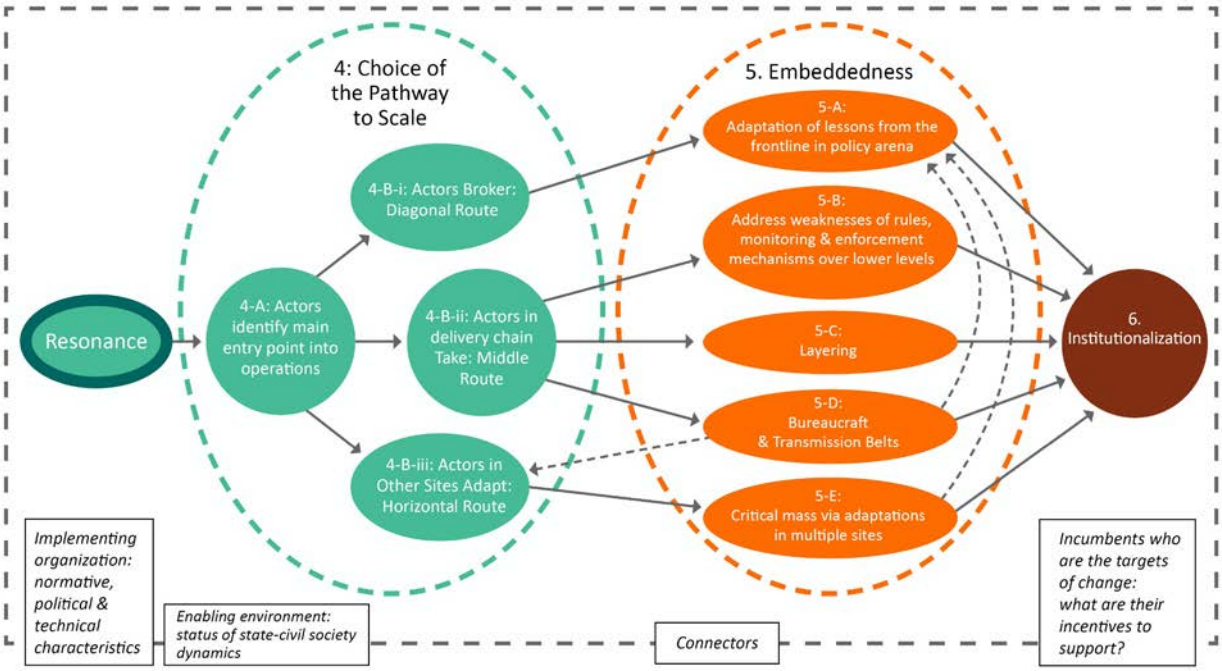
The idea of a resonance pathway emerged from collective insights from the social accountability practice of SALT research team members, including their experiences working with the Global Partnership for Social Accountability (GPSA), CARE, and World Vision, among other organizations. Although we suspect that the resonance pathway is the most promising one for scaling up in a large number of circumstances, the choice to pursue this pathway has often been unrecognized. A resonance pathway is less frequently specified in project strategies, theories of change, research and evaluation. This is perhaps, in part, because the main actors in this route are not the leading implementers of the interventions at the frontline (i.e. civil society groups). The lack of recognition for this pathway may also reflect a methodological failure to inquire into *why* public authorities decide to do things. Theorizing a pathway that relies on information about how public authorities make decisions, in which spaces, and why they decide to do certain things requires engaging with literature and evidence beyond civil society-focused social accountability.[25]

It is important to note that some of the steps in the resonance pathway may appear, at first glance, similar to the steps associated with other pathways. Theories about different pathways often use the same label to mean very different processes and mechanisms - as in the case of environmental and actor characteristics in Table 2. In a resonance pathway, stakeholders' goals and the steps they take as they navigate the pathway seek to work politically with the grain of the system to support the resolution of collective action problems in ways that fit and, potentially, contribute to changing service delivery over time. This political gamble, which often hinges on compromises between insiders in the public sector, is very different from those who advocate from outside the public sector for the wholesale adoption of universal norms through confrontation or best practices, as associated with other pathways. The assumption underlying the resonance pathway is not that norms or technical knowledge do not matter to decision making. Rather, assumptions that underlie the resonance pathway include that decision-makers often rely on their understanding of different pieces of knowledge and evidence, including different ideas and norms, to make pragmatic judgments. These judgements are based on assumptions about the reasons,

[25] For example, Greenghlagh, *et al.* (2004 and 2017)'s review of scalability of innovations in the health sector provide relevant cues. It can help understand the, often lengthy, processes by which agents of change' influence potential adopters' perceptions and negotiations to take complex social accountability innovations to scale.

explanations and causes for social accountability and how it can be scaled up with in their own context. These assumptions are often not sufficiently addressed theories of change.[26]

Figure 6: Unpacking Resonance and Embeddedness (indicative theory of change)



The idea of a resonance pathway challenges traditional conceptualizations of advocacy campaigns and the roles and approaches often assumed for civil society actors in resistance pathways (also see consistent findings about CARE's track record in Aston, 2019). The review of the literature and the evolution of practice calls for an alternative theorization that puts the spotlight on the resonance pathway as a potentially fruitful and under-assessed means to integrate social accountability insights into public sector reforms. The related focus of the theory of change on connectors and the quality of connection in the discussion of key actors' and environmental characteristics in Table 2 further reinforces the importance of this step.

Phase 4.A kick starts the path when actors, often working within the public sector or advising the public sector, perceive that social accountability contributes to solving problems relevant to their institutional goals. In this case, they will support scaling insights from social accountability from the frontline through the opportunities available to them in their context.

A small number of public officials and their advisors can help overcome challenges that civil society advocacy often faces when attempting to influence the policy arena with insights from the frontline. In fact, in this pathway, civil society organizations may have to take on different roles. Rather than acting as external monitors, campaigners, or implementers, civil society organizations may need to be facilitators. For example, in the embedded social accountability

[26] Pritchett (2021) articulates the rationale for assuming this decision-making process instead of the assumptions underlying advocacy for decision-makers only considering RCT-based evidence.

for a family planning model in Malawi's Ntcheu district, the district health management team had to take ownership and receive (modest) support from civil society, including capacity building in effective facilitation. CARE took on a more limited engagement, providing only targeted technical assistance and capacity building (CARE, 2020). This differed significantly for their previous intervention of resource-intensive direct implementation of a particular scorecard model.

Different insiders have different opportunities available to them. The diagonal, middle, and horizontal routes (**Phase 4 - B - i, ii, iii**) portray some of those unique opportunities and the factors that facilitate or hinder each path. These non-exclusive pathways are consistent with literature on health reforms that finds that they 'play out across different parts of government and different tiers of the public administration (Kelsall, Hart and Laws, 2016: 6, also WHO, 2010; Fox and Reich, 2015; Falleti, 2013).'

Phases 4 -B - i and 5 - A: Actors with access to the frontline and the policy arena (e.g. development partners, political leaders, academics, consultants, international or national non-governmental groups) use their roles as conveners, brokers, and advisors and researchers, among others, to support uptake in the policy arena. Uptake may be on a trial basis, confined to some areas and not others or to some parts of the delivery chain and not others. For example, an evaluation of a Save the Children Georgia and Civitas Georgia project focused on pre-school education found that the World Bank, which had supported the project, was picking up lessons. These lessons informed the conceptualization of a sectoral operation to support the government (ECORYS, 2020) - an insight from projects in Mongolia (Ali, 2019) and elsewhere that informs the GPSA's theory of action (Guerzovich *et al.* 2020). Ball and Westhorp (2018) and Long and Panday (2020) document the role of World Vision as a facilitator and connector in Indonesia and Bangladesh. Green (2016, 2017) makes a similar argument about the role of Oxfam in Tajikistan's water and sanitation sector. In the Philippines, Sidel and Faustino (2019) found that Coalitions for Change achievements in turning localized interventions into scalable and sustainable results. This happened when savvy connectors changing the inside game through which reforms are embedded into public sector practices and procedures. This tactic is in contrast to other social accountability initiatives that have opted for campaigns and making demands for policy change from the outside. These findings about pathways to change, as Green (2020) puts it, can be "uncomfortable/heretical" for many proponents of empowerment-led approaches to change.

Phase 4 - B – ii. identifies the middle routes through which stakeholders at the managerial and governance levels of health delivery chains use their mandates, social and political capital, and other resources to tap into lessons from the frontline to inform action in other points of the delivery chain.

Phase 5- B: Insights from social accountability in one site might be helpful to addressing problems in others which may partly account for spillovers with officials in neighboring locations requesting civil society groups for their "own" social accountability processes. These are, often, the results of non-linear processes and sustained by formal partnerships including with civil society or through recurring external mobilization and monitoring to prevent

reversals. In Bangladesh, World Vision's social accountability intervention catalyzed collective action at the local level, and informed regional and national officials which incentivized them to think creatively to find new ways to solve persistent policy design and implementation problems beyond those sites (Long and Panday, 2020).

Phase 5 -C: The space for influencing public sector decisions often exists in the managerial and governance levels (Levy and Walton, 2013; Levy, 2014), including in partly decentralized health delivery chains. These spaces can allow insiders to take actions to introduce lessons, processes, data and/or value from social accountability in other sites into the policy-making process. This can happen gradually and be disguised within the system, rather than triggering a backlash from more skeptical insiders that coexist with insider-led change efforts. These altered rules and ways do not necessarily overturn older rules, and can be placed on top of the older rules in layers, gradually changing the way a health delivery system works in practice (Thelen, 2004, Mahoney and Thelen, 2010).[27] For example, the technical officials from the Ministry of Health in Ghana engaged with the development of citizen alternative budgets. They helped citizens to understand sectoral budget programs and set realistic parameters for what citizens could ask for. When the CSOs presented their priorities to the Ministry and Parliament based on this guidance, the insiders had already spoken to their units in the Ministry, which paved the way for more effective advocacy in Parliament (Mills, 2019).

Phase 5- D: Bureaucrats and managers can also use their agency to navigate the system and exact responsiveness (Joshi and McCluskey, 2018). They can use their agency to strategically export ideas and interpretations about the value of social accountability from the frontline to the policy arena, leveraging the credibility that they have as insiders (Masaki *et al.* 2017; Carstensen and Röper, 2019). For example, an evaluation of social accountability for health implemented by CORDAID in DRC found that project managers who left the original sites of action to continue their career either in a different organization or in a different unit of CORDAID DRC are acting as connectors. They are spreading the ideas and lessons learned on the frontline and embedding them in other organizations. This includes using those ideas in the design of large UNICEF, World Bank, and USAID programs, as well as working with provincial health authorities (Falisse *et al.* 2019).

The last pathway to scale, **Phase 4 - A- iii**, is horizontal - as interventions in one locality inspire and encourage emulation from stakeholders in other localities. Public officials in other frontline locations learn insights from other social accountability action thanks to networks, spillover effects and/or advocacy actions. While, for the external observer, these appropriations may look like "best practice" replication, a resonance route accepts compromises and trade-offs associated with contextual characteristics, organizational capacities and collective action - rather than benchmarking and standardizing from the top-down or the outside in.[28] Implementing and testing adaptations elsewhere may or may not

[27] Falleti (2010) identifies and illustrates this process at work in the transformation of the Brazilian health system, from experimentation at the local level to scale up, through the subversion of rules from within the public sector.

[28] Guerzovich and Schommer (2017) identify these two dynamics at work in the multiplication of social observatories in Brazilian municipalities arguing that different pathways produced different results.

help attain a critical mass and, in so doing, support embeddedness in other levels of the (**Phase 5-E**).

As Figure 5 illustrates, these options can be hybrid, connected at different stages in the chain. Insiders negotiate, often in an iterative fashion, the internalization and adaptation of insights and principles from the frontline, and can secure their reapplication in government practices and procedures. This means that the initial civil society-led intervention's design characteristics may be less relevant than political and technical processes up the chain. According to Greenghalgh *et al.* (2004) this is a feature of intervention scale up in service delivery sectors, beyond social accountability: "the attributes (of an innovation) are neither stable features of the innovation nor sure determinants of their adoption or assimilation. Rather, it is the interaction of the intended adopter(s), and a particular context that determines the adoption rate" (Greenhalgh *et al.* 2004: 598).

Embeddedness and institutionalization are generally the product of actors' direct intentions, even if the specifics of scale up sometimes reflect *unintended* consequences that do not match any particular actors' preferred choice and framing of that choice (Thelen, 2004; Mahoney and Thelen, 2010). The process of enabling embeddedness also means that there will be a progressive transition of ownership as well as likely adaptation of practices along the way. In other words, the approach to procedural sustainability offers more pathways to scaling and financial support focuses on *function* ("identifying the systems and functions which need to be in place in order to support an ongoing process of state-citizen interaction around a particular problem or problem area") rather than *form* (a project and its owner, a technology, a tool, a standard) (Integrity Action, 2020). For many CSOs or NGOs, scaling-up a tested innovation might mean trade-offs and compromising or losing features that they value: from reduced margin of maneuver regarding the design of the intervention, to funding, to branding of the social accountability process. Hence, the resonance pathway is a process of "indigenous institutionalization" in which copying and borrowing often involves careful adaptation, along with greater local legitimacy and appropriateness (Leftwich, 2009) - and less control by the proponents of reforms (Integrity Action, 2020).

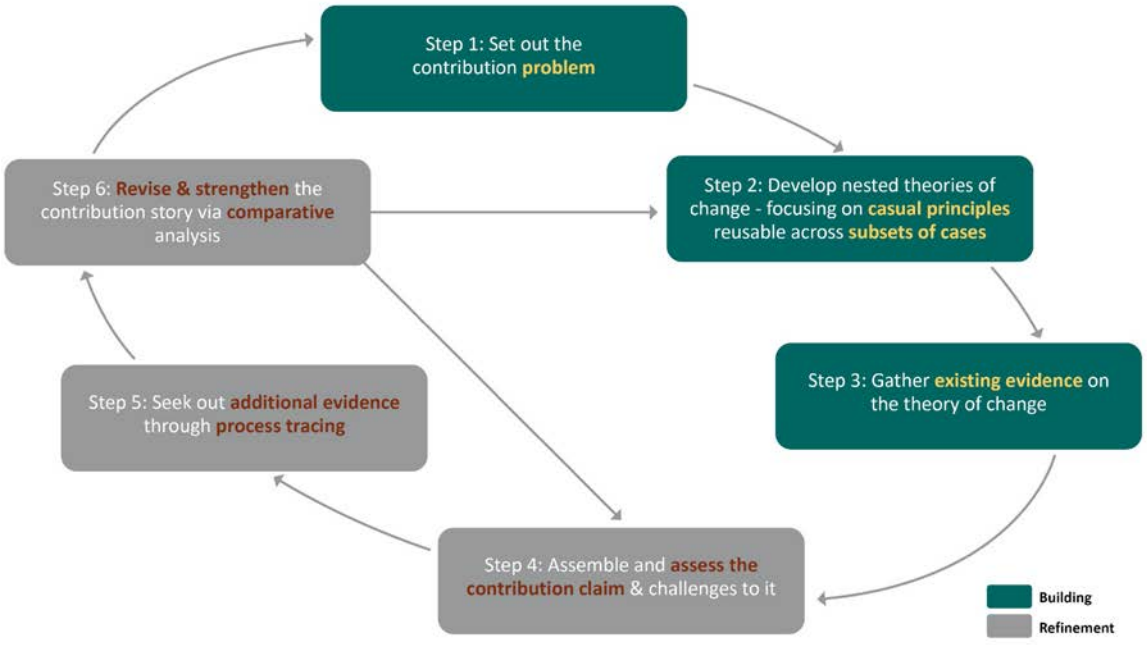
The concrete steps of the process and the associated goals, therefore, can be distinguished from other pathways in which outsiders, with support of inside champions, work to impose the wholesale adoption of first best proposals through advocacy at the highest political levels. Those processes and objectives are often associated with the resistance pathway, and are framed in terms of universal norms - with outsiders playing a necessary, if not sufficient, role to advance these norms through confrontation and agenda setting. They are also associated with the technocratic replication of best practices often associated with the alternative pathway.

VI. Further details of the project's technical design

This research project design paper reflects the initial steps of a process to build and refine middle-level theory through an iterative, abductive process that challenges and reconsiders blind spots in the theories of change used in social accountability evaluation, rather than a fully developed framework ready to be tested. It builds on the conceptual, normative, and methodological foundations of international practice. It does not challenge their soundness, because the goal is to theorize how and under what conditions these social accountability practices scale in their own terms or not.

This paper set out the contribution problem (paradox of mixed evidence and increasing uptake - Step 1 in Figure 7 below).

Figure 7: Iterating Theory Building and Refinement



Then, it illustrated how the portfolio of social accountability interventions is expected to work by developing a nested theory of change (Step 2, see Sections II and III above). In this nested theory of change, the note specifies 3 pathways to scale, including their different causal principles and critical boundary conditions. This last step is important to understand under which conditions the resonance pathway hypothesized to be at work in a limited range of cases will be likely to be more effective than plausible alternative explanations. This allows us to identify evidence which would help support or reject hypotheses with greater confidence through causal inference (Collier, 2011). A related contribution is to enable better understanding of what the routes to scale in the three pathways look like - between abstract,

universal generalizations and idiosyncratic descriptions - and when and where they might be a smart bet.

Step 3, the theory building exercise so far has drawn on existing evidence in the field, including the body of literature about social accountability and secondary literature from other relevant fields,[29] a consultation of more than 40 stakeholders in the field, and tacit knowledge about the trajectory of the 41 collaborative social accountability projects in the GPSA's portfolio, as well as hundreds of projects implemented by World Vision, CARE, and other civil society groups in more than 50 countries.

During the 12-month CEDIL-funded project, the team will use the aforementioned methods to refine its nested theory of change. An important step in this process was to assemble and assess the contribution claim and challenges to it, mainly by looking at the evidence available for competing claims in relation to social and formal sanctions, which lie at the heart of a resistance pathway to scale (see Aston and Zimmer Santos, forthcoming).

In the coming months, the team will collect and analyze data from a small number of carefully selected cases using a combination of process tracing and comparative case study methods (Step 5). Within each case, we will explore the presence of factors, contextual conditions, and causal mechanisms predicted by the theory, along with alternative explanations to draw valid causal inferences. By exploring within- and cross-country country variation using functional equivalences,[30] we can provide analyses of scale up processes in different policy contexts, while holding some idiosyncratic and country-specific characteristics constant. For this reason, case selection criteria is critical to strengthen the validity of this small-n study (Mahoney, 2000; Brady *et al.* 2004; White and Phillips, 2012) while adequately performing the function of supporting the refinement of the theory.

The nested theory of change theorizes the moderators that might be at work at the critical juncture where agents can opt for resonance or the other alternative sub-theories of change. This key insight, the in-depth engagement of the literature in preparation of this paper, as well as the uncertainties, increased costs for, and restrictions for in-person data collection due to the COVID-19 pandemic inform a second viable strategy for theory refinement, including case selection and data collection. With the pandemic preventing travel in 2020, the team was able to do a wider literature review which has allowed us not only to elaborate new theory, but to look in greater depth and tension points between potentially rival theories in particular contexts. Of particular interest is the finding of strategic pivots in a number of cases, which show how what are currently presented as rival theories may potentially interact over time as

[29] This includes secondary research about the institutionalization of participatory processes in the health sector and the broader literature about institutional change, policy diffusion and, thinking and working politically in development, governance and development, fragility conflict and violence, among others also helped specify the theoretical framework thus far (The World Bank, 2017, Levy and Walton 2013; Levy, 2014; Carpenter, 2001; Andrews, Woolcock and Pritchett, 2017; Honig, 2018; McDonnell, 2019; Pierson, 2004; Mahoney and Thelen, 2010; Falleti, 2010;, Rogers, 1995; Falleti and Lynch, 2009).

[30] More generally, on this methodological issue see Przeworski and Teune (1970); Locke and Thelen (1995).

civil society actors pivot strategy at critical junctures (Aston, 2015; Larsen, 2015; Poli, Guerzovich and Fokkelman, 2020).

The nested theory of change presented in the document introduces a key critical juncture when changemakers can opt for 3 alternative pathways to scale. This critical juncture forces researchers and evaluators to move beyond a zero-sum debate between alternative sub-theories of change and explore the conditions under which each pathway might offer a more promising course of action. We consider that this analytical shift is a critical one to produce transferrable evidence about social accountability's scalability, addressing the puzzle that informs this project and closing the gap between theory and practice. This critical juncture emerged as part of the theory building process and presented the team with the need to reconsider the best combination of cases to refine the theory to explicitly look for "critical cases" that can enable us to learn about the limits and boundaries of alternative pathways.

Some of the proposed research cases were selected specifically to refine the resonance sub-theory of change for three reasons. First, at the frontline, the social accountability interventions implemented in countries are typical ones in the field, and thus useful for engaging the literature and mixed results and building theory through process-tracing (Beach and Pederson, 2019). Second, in each country, the team had already identified reforms processes under way that would enable the research to learn about the causal mechanisms and moderators that may enable, at least, partial embeddedness of lessons from the frontline at scale. Third, key stakeholders in government, civil society, and the World Bank have expressed interest in the exercise, which facilitates its feasibility within a limited period of time and limited resources.

The research proposes to include a total of 8 cases across 6 countries. The unit of analysis is a particular type of policies and programs implemented by public sector actors which aim to strengthen health systems through lessons from social accountability through functional equivalent instruments prioritized at country level. The cases under study are purposefully selected policies or programs which represent the aforementioned tension points in the literature. The focus of each policy or program arena may differ between cases depending on the nature of scale up. We propose to organize the cases in 3 tiers, which represent different degrees of data collection and intensity of effort.

The first tier will be assessed through virtual interviews and document revision. Some members of the team conducted preliminary research in Indonesia in July 2019, and the research team has a broad range of contacts that would facilitate virtual data collection through World Vision (Indonesia and Cambodia). The cases selected are presented in table 3 below (see also *Methodological Design and Interview Guide*):

Table 3. Tier 1 Cases

Country	Cases
Indonesia	• National Strategy to Accelerate Stunting Prevention (StraNas Stunting) (2014) • Implementation of the Village Strengthening Law (2014-)
Cambodia	• Implementation of Social Accountability Framework (I-SAF) (2016-)

In general, social accountability is "lumpy," coming from very few countries. Indonesia is one context which is frequently cited in evidence reviews (see Tsai *et al.* 2019). Moreover, it represents a crucial context to assess the role of rival causal pathways. It is one context in which there is some evidence that the replication of best practice, resistance, and resonance pathways may be at work in interrelated programs in the health, nutrition and village strengthening sectors (see Ball and Westhorp, 2018; Gaduh *et al.* 2020). Our preliminary research suggests there may be some relationship to the uptake of social accountability approaches from Wahana Visi and the GPSA programming. So, it is also possible to draw on some tacit knowledge of the process. However, conscious of potential positive and confirmation biases, we have deliberately chosen researchers in the team without any personal connections to the programming for conducting primary research.

On the other hand, Cambodia is a country context which rarely appears in evidence reviews in the social accountability sector. Yet, it is a context in which there are some social accountability operations of substantial scale and longevity. It is therefore apt for further exploration. In this case, the research team is also able to leverage tacit knowledge of the framework from World Vision, CARE, and World Bank experience. Yet, knowledge gaps remain regarding the process adoption and scale up of the framework with donors and the government. As in the Indonesia cases, we have deliberately chosen researchers in the team without any personal connections to the programming for conducting primary research.

The second tier is similar to tier 1, but takes particular account of Fragile and Conflict-affected State (FCAS) contexts and the type of outcomes which are likely to be achieved in such contexts. Focusing on partial, sub-optimal results as 'positive deviants' is important because previous research argues that 'clear and big accountability gains' are deemed 'unlikely' and that 'backlash' is judged to be 'inevitable (Joshi, 2019: 2-6).' And, yet, this assessment relies on a particular definition of sanction-based accountability and potentially unrealistic expectations of progress, whether in fragile or more stable settings. Positive deviance, within typical countries, can be useful in theory building exercises (Roche, 2015). To this end, we propose to consider the scale up of social accountability mechanisms linked to the Sanitary Development Committees in the DRC.

Table 4. Tier 2 Case

Country	Cases
DRC	• Health Area Development Committees - CODESA (*Comité de Développement Sanitaire*)

The DRC provides a good example to explore the potential and plausibility of the resonance pathway for uptake of social accountability in the health sector in fragile and conflict affected contexts. In general, DRC is argued to be a typical fragile state, given its weak capacity to provide services and security, limited political will to introduce reforms, with strong neo-patrimonial structures and many different actors competing with the state for legitimacy and authority (Lindström, 2019).

In particular, the chosen case in the DRC will allow the research to look into scale up at sub-national level (provinces and health zones), which may be a more salient level than national level in a context such as DRC. It also allows the research to inquire into linkages with international aid actors and how innovations in social accountability mechanisms may influence the set-up of other mechanisms in other sectors such as education (see Falisse *et al.* 2019). As this case is related to findings from a GPSA project, this allows us to identify relevant stakeholders who can help shed further light on potential scale up of the aforementioned mechanisms. We, however, propose to employ the same rule as those in tier 1 in order to ensure researchers' independence.

The third tier represents subsidiary case studies. These are designed, in particular, to probe further on potentially rival explanations and boundary conditions. These are cases where the research team has identified: a) critical junctures when local changemakers considered and/or took alternative pathways to scale; b) the team has access to relevant secondary literature, including published and unpublished evaluations and project documents and/or research, that provides insights into those critical junctures; c) the team has assessed that it will be able to address gaps in available information through email correspondence in each case. None of the tier 3 cases will gather additional primary data, beyond correspondence.

Table 5. Tier 3 Cases

Country	Cases
Uganda	• Health financing in National Budget Framework (2011 - 2018)
Malawi	• Community engagement in the National Community Health Strategy (2017 - 2022)
Peru	• Financing and implementation of National Nutrition Strategy - CRECER (2007-2014) • Citizen monitoring linked to the Health Reform (2009 - 2014)

The case in Uganda is a canonical example in the sector for a resistance pathway, manifested in the Human Rights for Health (HRH) campaign. Yet, it also illustrates important critical junctures between at least two of the pathways proposed for further study in this research. We are able to draw on peer reviewed research and grey literature on the campaign (Larsen, 2015; Walker, 2017; Templeton Dunn *et al*. 2017), as well as tacit knowledge of the campaign itself, as World Vision co-coordinated the campaign.

In the case of Malawi, similarly, existing research on the case suggests a tension point between at least two of the pathways and potential pivots between these. We are also able to draw on tacit knowledge of the programming and efforts to scale up scorecards at CARE, where two of the research team worked previously, including data from monitoring and evaluation systems (Gullo *et al.* 2017; CARE, 2020).

One of the Peruvian cases, CRECER, is also a canonical example in the sector for the potential role of campaigns. Most notably, the widely cited influence of the Child Malnutrition Initiative (CNI). We can draw on peer reviewed research and independent evaluations (Acosta, 2011; Yosef and Goulden, 2016; Aston, 2019). The case is also of particular importance to this research because there was a specific exchange between Peru and Indonesia. As the CNI was coordinated by CARE, we are also able to draw on the tacit knowledge of our network involved in the campaign and the tracking of commitments over time in the CRECER strategy over time. In the other Peru case, CARE is also a common link, and yet it is a case in which more than a single pathway was likely in play over the medium term. We can draw on peer reviewed research, grey literature, independent evaluations (Ackerman, 2005; Vargas, 2013; Aston, 2015; Samuel *et al.* 2020), as well as tacit knowledge of our network.

For all three tiers, the research team will use process tracing as the primary methodology. Evaluation experts (Stern *et al.* 2012), researchers focused on the broader institutional changes social accountability seeks to inform (e.g. Mayka, 2019; Falleti and Cunial, 2018) and proponents for prominent social accountability theories (Joshi, 2014; Fox, 2015; Fox, 2020) have all identified the potential of process tracing for its strong depictions of a program's

theory, focusing on processes of change within cases, which may uncover causal mechanisms. To realize this potential the team will, first, make predictions about what empirical evidence would be left by an intervention if the hypothesis were true, considering evidence we would both expect and hope to find, and assess the "probative value" of this evidence. To do this, we will employ an informal Bayesian updating approach, which is also a recommended method for causal inference in advocacy efforts (Neave *et al.* 2017, also Box 3 below).

Box. 3: Bayesian Reasoning

In recent years, there has been increased interest in moving from analogies to the formal application of Bayesian logic to process tracing (Bennett, 2008; Bennett, 2014; Schmitt and Beach, 2015; Humphreys and Jacobs 2015; Mahoney, 2016; Befani and Stedman-Bryce, 2016; Fairfield and Charman, 2017, 2018; Beach and Pedersen, 2019). Fairfield and Charman (2017) view the turn to Bayesianism as a watershed for in-depth, small-n research. Bayesian reasoning is a means of updating one's views regarding which hypothesis best explains the phenomena or outcomes of interest as we learn additional information (Fairfield and Charman, 2018: 6). Key to that updating process and our level of confidence is how unique specific pieces of empirical evidence are to a particular hypothesis (Beach and Pedersen, 2019). Relatively few efforts have been made to include Bayesian updating in social accountability or advocacy evaluations. Wadeson, Monzani and Aston (2020) compared different types of process tracing for theory-testing, including full Bayesian updating through 'contribution tracing' as well as more agile approaches such as 'contribution rubrics.' The study found that while full Bayesian updating can be beneficial for theory testing at the micro level, it has limits when working with loose theories of change (Wadeson, Monzani and Aston, 2020), and for theory-building forms of process tracing. As such, more informal Bayesian reasoning is recommended for this study.

Second, the research will conduct in-depth empirical observation and enquiry, including interviews with civil society actors working on a project/sector, along with public officials, development partners and other relevant actors. Data collection, including consent, and management will be in compliance with the standards and procedures of the University of Santa Catarina Brazil (see *Methodological Design and Interview Guide*).

Each case study will include an assessment of the perceived value of social accountability to wider reform efforts and will describe in detail the specific mechanisms through which social accountability was possible to scale. This will also include consideration of where this potential may be derailed. Having identified that social accountability is perceived to and/or adds value at scale – and how – the research will explore what enabled the incorporation of social accountability at scale. To do so, the team will: a) construct a timeline for each case study with the sequence of steps of policy making and implementation, including the steps involved in the incorporation and adaptation of social accountability (or its lessons and/or component parts), b) identify key events in the timeline in the pathway towards the incorporation and adaptation of social accountability; and c) drill down into the details of how and why each critical juncture unfolded the way that it did. This last step will enable the team to answer questions vis-à-vis social accountability's incorporation at scale. For each critical juncture, the team will identify the range of stakeholders involved in decision processes vis-à-vis incorporation of social accountability (or its lessons and/or component parts), with careful assessment of how and why the facilitators of social accountability's adoption and adaptation were able to win support for its inclusion and scaled-up.

Third, the team will assess whether we can trust the evidence found for each part of the hypothesized theory of change in the identified cases (figure 7). The goal of this assessment is to update the study's confidence in the hypothesis identified before fieldwork, and possibly develop alternative hypotheses, particularly in response to downgraded confidence in the pre-fieldwork hypothesis developed.[31] Based on these three features, one can make a judgement regarding whether hypothesized mechanisms are present or not (Beach and Pedersen, 2019: 4, 178). Causal process observations (CPOs) or "diagnostic evidence" used to support theories should be linked to empirical fingerprints (George and Bennett, 2005: 6; Brady and Collier, 2011; Bennett and Checkel, 2014). In the social accountability field, these fingerprints are not often explicitly demonstrated in evidence. Consequently, studies and evaluations commonly fall short of providing evidence that may clearly confirm, reject or refine theories of change and action to support decision-making, learning, and accountability over time (see Joshi, 2011; Anderson, Fox, and Gaventa, 2020). This is an important gap to fill between evaluation and practice.[32]

[31] See *Methodological Design and Interview Guide* for further details of question probes for rival and/or complementary pathways.

[32] For interested readers, Annex 1 discusses in greater detail common limitations of process tracing and the study's proposed responses to these.

Process tracing will be complemented with comparative historical analysis to strengthen the theory building and inferential goals of this study (Mahoney and Rueschemeyer, 2003). Unlike the majority of studies in the social accountability field, we consider that causal mechanisms can also allow us to advance meaningful bounded comparisons across cases (Elster, 1989; Bates, 1998; Hall, 2003; Hedström, 2008). They can help move the focus of social accountability research and evaluation beyond the results of individual projects, to how and why they may be integrated within the political economy of policy-making processes in the health sector and to what effects.

A comparative synthesis of these case studies will extract lessons from an additional set of secondary cases (tier 3). The goal of this exercise is to help further refine the theoretical propositions about the causal mechanisms of the proposed theory of change, along with alternative explanations and their boundary conditions (Step 6). The dialogue of the case study findings with the literature, as synthesized and interpreted here, will help produce a series of hypotheses and a research design, with the future goal of testing the theory in a broader range of cases, including from other relevant sectors such as education. As this process unfolds, it is informed by the development of guidance for grant partners and evaluators of projects funded by the Global Partnership for Social Accountability, including guidance for monitoring, evaluation, reporting and learning of questions about scalability and sustainability.

Over time, as new evidence emerges and the theory is refined (loop connecting Step 6 to Step 4 in Figure 7), a more radical revision of the theory could be in order (loop connecting Step 6 to Step 2 in Figure 7). To this effect, the team will prepare a research proposal mapping how the methodological innovation might be tested in other service delivery chains, including but not limited to the health sector. As there is a gap between theory development and evaluation practice in social accountability, guidance for using process tracing abductively can inform practitioners and other researchers and evaluators in the development and refinement of individual projects' and programs' theories of change, as well as in better understanding how their work in concrete contexts maps onto portfolios and the broader field, so they can support evidenced-based strengthening of practice field-wide.

Bibliography

Ackerman J M (2005) Human rights and social accountability. Participation and Civic Engagement, Social Development Department, Environmentally and Socially Sustainable Development Network. Washington, DC: World Bank. Available at: http://www.accountabilityindia.in/sites/default/files/document-library/102_1237790940.pdf

Acosta A (2011) Analysing Success in the Fight against Malnutrition in Peru. IDS Working Paper 367, IDS, Brighton.

Ali R (2019) Implementation Completion and Results Report (ICR) Document -Transparency and Accountability in Mongolian Education (TAME) - P150842. Washington, DC: World Bank Group. Available at: http://documents.worldbank.org/curated/en/527361562846501495/Implementation-Completion-and-Results-Re-port-ICR-Document-Transparency-and-Accountability-in-Mongolian-Education-TAME-P150842

Anderson C, Fox J and Gaventa J (2020) How Do Donor-led Empowerment and Accountability Activities Take Scale into Account? Evidence from DFID Programmes in Contexts of Fragility, Conflict and Violence. IDS Working Paper 536, IDS, Brighton.

Andrews M, Woolcock M and Pritchett L (2017) Building state capability: Evidence, analysis, action . Oxford, UK: Oxford University Press.

Arkedis J, Creighton J, Dixit A, Fung A, Kosack S, Levy D and Tolmie C (2019) Can Transparency and Accountability Programs Improve Health? Experimental Evidence from Indonesia and Tanzania. HKS Faculty Research Working Paper Series RWP19-020, Harvard Kennedy School, Cambridge, MA. Available at: https://www.hks.harvard.edu/publications/can-transparency-and-accountability-programs-improve-health-experimental-evidence

Aston T (2015) Citizen Monitoring to Defend Maternal Health Rights in Peru. Learning & Policy Series Briefing Paper, 6. Available at: https://insights.careinternational.org.uk/media/k2/attachments/CARE_Citizen-monitoring-in-Peru_2015.pdf

Aston T (2016) What has CARE Learned from Political Economy Analysis. Global Partnership for Social Accountability. Available at: https://gpsaknowledge.org/wp-content/uploads/2016/01/What-has-CARE-has-learned-from-political-economy-1.pdf

Aston T (2019) CARE International Advocacy and Influencing: A Review of Pathways to Success. Available at: http://www.careevaluations.org/wp-content/uploads/CARE-Advocacy-and-Influencing.pdf

Aston T (2021) Nuggets, Smoking Guns, and the Wild West of Evaluation, available at https://www.cdimpact.org/events/nuggets-smoking-guns-and-wild-west-evaluation

Aston T and Zimmer Santos, G F (forthcoming). Social Accountability and Sanctions: A Review.

Ball D and Westhorp G (2018) Citizen Voice and Action for Government Accountability and Improved Services: Maternal, Newborn, Infant and Child Health Services. Final Evaluation Report. Community Partners Ltd., Jakarta, Indonesia. Available at: https://gpsaknowledge.org/wp-content/uploads/GPSA-Indo-nesia-CVA-Project-Final-Evaluation-Report-2018.pdf

Banerjee A, Duflo E and Glennerster R (2007) Putting a band-aid on a corpse: incentives for nurses in the Indian public health care system. Journal of the European Economic Association 6: 487–500.

Batley R and Mcloughlin C (2015) The politics of public services: A service characteristics approach. World Development 74: 275-285.

Bates R H (1998) Analytic narratives. Princeton, N.J: Princeton University Press.

Beach D (2016) It's all about mechanisms–what process-tracing case studies should be tracing. New Political Economy 21(5): 463-472.

Beach D and Pedersen R B (2019) Process-tracing methods: Foundations and guidelines. Ann Arbor, Michigan: University of Michigan Press.

Behn R D (2001) Rethinking Democratic Accountability. Washington, DC: Brookings Institution Press.

Bennett A and Checkel J (Eds.) (2014) Process Tracing in the Social Sciences. From Metaphor to Analytic Tool. New York: Cambridge University Press.

Björkman M and Svensson J (2009) Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda. Quarterly Journal of Economics 124(2): 735–69.

Booth D (2012) Development as a collective action problem: Addressing the real challenge of African governance, Africa power and politics programme. Synthesis Report, Overseas Development Institute, London. Available at: https://www.odi.org/sites/odi.org.uk/files/resource-documents/appp-synthesis-report-development-as-a-collective-action-problem-david-booth-oct-2012_1.pdf

Bold T, Kimenyi M, Mwabu G, Ng'ang'a A and Sandefur J (2013) Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education. Working Paper 321. Center for Global Development, Washington, DC. Available at: https://www.cgdev.org/publication/scaling-what-works-experimental-evidence-external-validity-kenyan-education-working

Boydell V and Keesbury J (2014) Social accountability: What are the lessons for improving family planning and reproductive health programs? review of the literature. Working paper. Population Council, Evidence Project, Washington, DC. Available at: https://knowledgecommons.popcouncil.org/cgi/viewcontent.cgi?article=1942&context=departments_sbsr-rh

Brady H E, Collier D and Seawright J (2004) Refocusing the discussion of methodology. In Brady H E and Collier D (Eds.) Rethinking social inquiry: Diverse tools, shared standards (pp. 3-20). Lanham, Maryland: Rowman & Littlefield Publishers.

Brady H E and Collier D (Eds.) (2011) Rethinking social inquiry: Diverse tools, shared standards. Lanham, Maryland: Rowman & Littlefield Publishers.

Briggs X D S (2008) Democracy as problem solving: Civic capacity in communities across the globe. Cambridge, MA: MIT press.

Broome A, Homolar A and Kranke M (2017) Bad science: International organizations and the indirect power of global benchmarking. European Journal of International Relations 24(3): 514-539.

Bukenya B, Hickey S and King S (2012) Understanding the role of context in shaping social accountability interventions: towards an evidence-based approach. Manchester: Institute for Development Policy and Management, University of Manchester. Available at: https://www.researchgate.net/profile/Sophie_King4/publication/312085782_Understanding_the_role_of_context_in_shaping_social_accountability_interventions_towards_an_evidence-based_approach/links/586e75ee08aebf17d3a78715/Understanding-the-role-of-context-in-shaping-social-accountability-interventions-towards-an-evidence-based-approach.pdf

Cant S and Sarriot E (2019) Collective action for improving health services. Why the evidence will continue to be mixed and what to do about it. Available at: http://systems4health.blogspot.com/2019/11/collective-action-for-improving-health.html

CARE (2020) Government-led Community Score Card© for family planning services. Available at: https://insights.careinternational.org.uk/media/k2/attachments/CARE_Government-led-CSC-family-planning-services_July-2020.pdf

Carpenter D P (2001) The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies, 1862-1928 (Vol. 173). Princeton, NJ: Princeton University Press.

Carstensen M B and Röper N (2019) Invasion From Within: Ideas, Power, and the Transmission of Institutional Logics Between Policy Domains. Comparative Political Studies 52(9):1328-1363.

Cartwright N (2020) Using middle-level theory to improve programme and evaluation design. CEDIL Methods Brief. Oxford: CEDIL.

Chenoweth E and Stephan M (2012) Why Civil Resistance Works: The Strategic Logic of Non-violent Conflict. New York: Columbia University Press.

Collier D (2011) Understanding Process Tracing. Political Science and Politics 44(4): 23–30.

Cornwall A (2002) Locating Citizen Participation. IDS Bulletin 33(2): I-X.

Cornwall A and Coelho V S (Eds.) (2007) *Spaces for change? The politics of citizen participation in new democratic arenas* (Vol. 4). London: Zed Books.

Creighton J, Arkedis J, Fung A, Kosack S, Levy D and Tolomie C (2020) Insights from Transparency and Accountability Action Plans in Indonesia and Tanzania. Ash Center for Democratic Governance and Innovation, Harvard Kennedy School. Available at: https://www.healthynewbornnetwork.org/hnn-content/uploads/Insights-from-Transparency-and-Accountability-Action-Plans-in-Indonesia-and-Tanzania.pdf

Davey C, Hargreaves J, Hassan S, Cartwright N, Humphreys M, Masset E, Prost A, Gough D, Oliver S and Bonell C (2018) Designing Evaluations to Provide Evidence to Inform Action in New Settings. CEDIL Inception Paper No 2: London.

DFID (2014) Assessing the Strength of Evidence: How to Note, March 2014. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf

DFID (2019) Governance for Growth, Stability and Inclusive Development. DFID Position Paper. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/786751/Governance-Position-Paper2a.pdf

Doin G A, Dahmer J, Schommer P C and Spaniol E L (2012) Mobilização social e coprodução do controle: o que sinalizam os processos de construção da Lei da Ficha Limpa e da Rede Observatório Social do Brasil de Controle Social. Pensamento & Realidade 27(2): 56-79.

Duflo E, Hanna R and Ryan S P (2012) Incentives work: Getting teachers to come to school. American Economic Review 102(4): 1241-78.

Duflo E and Kremer M (2015) Which Innovations Reach More than 100,00 or One Million People? Evidence from the Development Innovation Ventures Portfolio. Available at: https://www.usaid.gov/sites/default/files/documents/15396/Which_Innovations_Reach_More_Than_100K_or_More_Than_1M_People_20150403_MK525PM.pdf

ECORYS (2020) Final External Evaluation of the Improving Preschool Education in Georgia through Social Accountability Processes Project. Final Report (DRAFT), Ecorys Polska.

Edwards D, Lopez Franco E, and Shutt C (2020) Open Up Contracting: End-Term Evaluation of the Citizen Agency Consortium Open Up Contracting Program, Open Contracting.

Elster J (1989) Nuts and bolts for the social sciences. Cambridge, New York: Cambridge University Press.

e-Pact (2016) Macro Evaluation of DFID's Policy Frame for Empowerment and Accountability — Empowerment and Accountability Annual Technical Report 2016: What Works for Social Accountability. Final version. ePact, Oxford, UK. Available at: https://itad.com/wp-content/uploads/2017/06/EA-Macro-Evaluation-Technical-report-Dec16-FINAL.pdf

Falleti T G (2010) Infiltrating the State: The Evolution of Health Care Reforms in Brazil, 1964-1988. In Mahoney J and Thelen K (Eds.) Explaining institutional change: Ambiguity, agency, and power (38-62). Cambridge, UK: Cambridge University Press.

Falleti T G (2013) Decentralization in Time: A Process-Tracing Approach to Federal Dynamics of Change. In Arthur Benz and Jörg Broschek (eds.) Federal Dynamics: Continuity, Change and Varieties of Federalism. Oxford: Oxford University Press, 140-166.

Falleti T G and Cunial S L (2018) Participation in social policy: public health in comparative perspective. Cambridge, UK: Cambridge University Press.

Falleti T G and Lynch J F (2009) Context and causal mechanisms in political analysis. Comparative political studies 42(9): 1143-1166.

Falisse J, Mafuta E and Mulongo P (2019) Reinforcing Social Accountability in Health Services in Sud Kivu and Kongo Central Provinces. Final Evaluation of the GPSA/CODESA Project. Cordaid, Kinshasa, Democratic Republic of Congo. Available at: https://gpsaknowledge.org/knowledge-repository/reinforcing-social-accountability-in-health-services-in-sud-kivu-and-kongo-central-provinces/

Feruglio F (2017) Do more empowered citizens make more accountable states? Power and legitimacy in legal empowerment initiatives in Kenya and South Africa. Making All Voices Count Research Report, IDS, Brighton.

Fox J (2015) Social Accountability: What does the Evidence Really Say. World Development 72: 346-612.

Fox J (2016) Scaling accountability through vertically integrated civil society policy monitoring and advocacy. Working Paper, Making All Voices Count.

Fox J (2019) How Can a Rethink of Lessons from Field Experiments Inform Future Research in Transparency, Participation and Accountability? Evidence Matters (Blog). International Initiative for Impact Evaluation (3ie). Available at: https://www.3ieimpact.org/blogs/how-can-rethink-lessons-field-experiments-inform-future-research-transparency-participation

Fox J (2020) Contested terrain: International development projects and countervailing power for the excluded. World Development 133: 104978.

Fox A M and Reich M R (2015) The politics of universal health coverage in low- and middle-income countries: a framework for evaluation and action. Journal of Health Politics, Policy and Law 40(5): 1023–60.

Frisancho A and Goulden J (2008). “Rights-based Approaches to Improve People’s Health in Peru,” The Lancet, Vol. 372, December 13, 2008.

Fukuyama F (2004) State-building: governance and world order in the 21st century. New York: Cornell University Press.

Fung A and Wright E O (2003) Deepening democracy: Institutional innovations in empowered participatory governance (Vol. 4). London, New York:Verso.

Funnell S C and Rogers P J (2011) Purposeful program theory: effective use of theories of change and logic models. San Francisco: Jossey-Bass/Wiley.

Gaduh A, Pradhan M, Priebe J and Susanti D (2020) Scores, Camera, Action? Incentivizing Teachers in Remote Areas. RISE Working Paper Series 20/035.

Gaventa J and Barrett G (2010) So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement. IDS Working Paper 347, Institute of Development Studies, Brighton, UK. Available at: http://unpan1.un.org/intradoc/groups/public/documents/un-dpadm/unpan044380.pdf.

Gaventa J and McGee R (2010) Introduction: Making Change Happen - Citizen Action and National Policy Reform. In Gaventa J and McGee R (Eds.) Citizen Action and National Policy Reform: Making Change Happen. London: Zed Books.

Gbeleou S C and Schechter S (2020) How NGOs Can Work With Governments to Build Partnerships That Will Scale. Available at: https://ssir.org/articles/entry/how_ngos_can_work_with_governments_to_build_partnerships_that_will_scale#bio-footer

George A L and Bennett A (2005) Case Studies and Theory Development in the Social Sciences. Cambridge, MA: The MIT Press.

Grandvoinnet H, Aslam G and Raha S (2015) Opening the Black Box: The Contextual Drivers of Social Accountability. New Frontiers of Social Policy Series. Washington, DC: World Bank. Available at: http://documents1.worldbank.org/curated/es/235561468321545065/pdf/958090PUB0seri0PUBLIC09781464804816.pdf

Green D (2016) How change happens. Oxford, UK: Oxford University Press.

Green D (2017) Theories of Change for Promoting Empowerment and Accountability in Fragile and Conflict-Affected Settings. IDS Working Paper 499, IIDS, Brighton.

Green D (2020) Book Review: 'Thinking and Working Politically in Development'. Available at: https://oxfamblogs.org/fp2p/book-review-thinking-and-working-politically-in-development/

Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, Hinder S, Fahy N, Procter R and Shaw S (2017) Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. Journal of Medical Internet Research 19(11): e367.

Greenhalgh T, Robert G, Macfarlane F, Bate P and Kyriakidou O (2004) Diffusion of innovations in service organizations: systematic review and recommendations. The Milbank Quarterly 82(4): 581–629.

Guerzovich, M. F. (2020). Conquering our blind spots: Glamorizing the everyday politics of accountability (Part 1). Global Partnership for Social Accountability. Available at: https://gpsaknowledge.org/conquering-our-blind-spots-glamorizing-the-everyday-politics-of-accountability-part-1/

Guerzovich M F and Poli M (2014) Are We Ready for Strategic Social Accountability? Global Partnership for Social Accountability Series Dissemination Notes 1–8. World Bank, Washington, DC. Available at: https://www.thegpsa.org/related-resource/are-we-ready-strategic-social-accountability

Guerzovich M F and Poli M (2020) How Social Accountability Strengthens Cross-Sector Initiatives to Deliver Quality Health Services? Global Partnership for Social Accountability Note 17. World Bank, Washington, DC. Available at: https://gpsaknowledge.org/uncategorized/how-social-accountability-strengthens-cross-sector-initiatives-to-deliver-quality-health-services/

Guerzovich M F, Poli M and Fokkelman E (2020) The Learning Crisis and Its Solutions: Lessons from Social Accountability for Education. Global Partnership for Social Accountability Note 13. World Bank, Washington, DC. Available at: https://gpsaknowledge.org/knowledge-repository/the-learning-crisis-and-its-solu-tions-lessons-from-social-accountability-for-education/.

Guerzovich M F, Thindwa J, Jespersen A, Libresco B, Poli M and Fokkelman E (2020) Global Partnership for Social Accountability: Theory of Action. Global Partnership for Social Accountability, Washington, DC. Available at: https://www.thegpsa.org/Data/gpsa/files/field/documents/02.09.2020._gpsa_theory_of_action.pdf

Guerzovich M F and Schommer P C (2016) Four Ways in Which Social Accountability and Open Government Interventions Bridge the State and Society. Paper presented at the 12th ISTR Conference, Stockholm, Sweden, June 28 – July 1.

Guerzovich, M F and Schommer P C (2017) The Politics of Open Contracting for Urban Service Delivery: Brazilian Contexts, Strategies, and Learning. Available at: https://ssrn.com/abstract=3423180

Gugerty, M K and Karlan D (2018) The Goldilocks Challenge: Right-Fit Evidence for the Social Sector. New York, NY: Oxford University Press.

Gullo S, Galavotti C, Sebert Kuhlmann A and Msiska T (2017) Effects of a social accountability approach, CARE's Community Score Card, on reproductive health-related outcomes in Malawi: A cluster-randomized controlled evaluation. PLoS ONE 12(2): e0171316.

Gurung G, Derrett S, Hill P and Gauld R (2019) The role of social audit as a social accountability mechanism for strengthening governance and service delivery in the primary health care setting of Nepal: a qualitative study. Critical Public Health 1–12.

Hall P A (1986) Governing the Economy: The Politics of State Intervention in Britain and France. Oxford, UK: Oxford University Press.

Hall P A (2003) Aligning Ontology and Methodology in Comparative Politics. In Mahoney J and Rueschemyer D (Eds.) Comparative historical analysis in the social sciences. Cambridge, New York: Cambridge University Press.

Halloran B (2015) Strengthening Accountability Ecosystems: A Discussion Paper. Think Piece, TAI, Washington, DC. Available at: http://www.transparency-initiative.org/archive/wp-content/uploads/2015/11/Strengthening-Accountability-Ecosystems.pdf

Hay C (2016) Process tracing: a laudable aim or a high-tariff methodology? New political economy 21(5): 500-504.

Hedström P (2008) Studying Mechanisms to Strengthen Causal Inferences in Quantitative Research. In Box-Steffensmeier J M, Brady H E and Collier D (Eds.) The Oxford handbook of political methodology. Oxford, New York: Oxford University Press.

Hernández A, Ruano A L, Marchal B, San Sebastián M and Flores W (2017) Engaging with Complexity to Improve the Health of Indigenous People: A Call for the Use of Systems Thinking to Tackle Health Inequity. International Journal for Equity in Health 16(1): 26.

Hernández A, Hurtig A, Goicolea I, San Sebastián M, Jerez F, Hernández-Rodríguez F and Flores W (2020) Building collective power in citizen-led initiatives for health accountability in Guatemala: the role of networks. BMC Health Services Research 20:416.

Hickey S and Hossain N (Eds.) (2019) The Politics of Education in Developing Countries. From Schooling to Learning. Oxford, UK: Oxford University Press.

Honig D (2018) Navigation by Judgment: Why and When Top-Down Management of Foreign Aid Doesn't Work. Oxford: Oxford University Press.

Honig D and Pritchett L (2019) The Limits of Accounting-Based Accountability in Education (and Far Beyond): Why More Accounting Will Rarely Solve Accountability Problems. Working Paper 510, Center of Global Development. Available at: https://www.cgdev.org/publication/limits-accounting-based-accountability-education-and-far-beyond-why-more-accounting-will

Integrity Action (2020) Citizen-centred accountability: how can we make it last? Available at: https://integrityaction.org/media/12393/integrity-action-sustainability-research-briefing-note.pdf

Jacobstein D (2020) What is the Work?. USAID Learning Lab (blog), January 21. Available at: https://usaidlearninglab.org/lab-notes/what-work.

Joshi A (2011) Annex 1 Service Delivery: Review of Impact and Effectiveness of Transparency and Accountability Initiatives. Transparency and Accountability Initiative. Available at: http://www.transparency-initiative.org/archive/wp-content/uploads/2011/05/impacts_annex1_final1.pdf

Joshi A (2013) Do They Work? Assessing the Impact of Transparency and Accountability Initiatives in Service Delivery. Development Policy Review 31(1): s29–s48.

Joshi A (2014) Reading the local context: a causal chain approach to social accountability. IDS Bulletin 45(5): 23-35.

Joshi A (2017) Legal Empowerment and Social Accountability: Complementary Strategies Toward Rights-based Development in Health? World Development 99: 160–172.

Joshi A (2019) Introduction: Accountability Amidst Fragility, Conflict, and Violence: Learning from Recent Cases. IDS Bulletin 50(3): 1-16.

Knack S, Parks B C, Harutyunyan A and DiLorenzo M (2020) How Does the World Bank Influence the Development Policy Priorities of Low-Income and Lower-Middle Income Countries? Policy Research Working Paper 9225, World Bank Group, Washington. Available at: https://documents1.worldbank.org/curated/en/852751588260919707/pdf/How-Does-the-World-Bank-Influence-the-Development-Policy-Priorities-of-Low-Income-and-Lower-Middle-Income-Countries.pdf

Kelsall T, Hart T and Laws E (2016) Political settlements and pathways to universal health coverage. ODI Working Paper 432. Available at: https://www.odi.org/sites/odi.org.uk/files/resource-documents/10382.pdf

Khemani S (2007) Can Information Campaigns Overcome Political Obstacles to Serving the Poor? In Devarajan S and Widlund I (Eds.) The politics of service delivery in democracies - better access for the poor (pp.56-69). Expert Group on Development Issues (EGDI).

Kosack S and Fung A (2013a) Confrontation and Collaboration. Ash Center for Democratic Governance and Innovation, Harvard Kennedy School. Available at: https://medium.com/challenges-to-democracy/confrontation-and-collaboration-73174aff5c85

Kosack S and Fung A (2013b) Transparency and Accountability Interventions: Making Sense of the Evidence. Ash Center for Democratic Governance and Innovation, Harvard Kennedy School. Available at: https://medium.com/challenges-to-democracy/transparency-and-accountability-interventions-making-sense-of-the-evidence-8efe2be6268d

Kosack S and Fung A (2014) Does Transparency Improve Governance? Annual Review of Political Science 17(1): 65-87.

Kosack S, Bridgman G, Creighton J, Tolmie C and Fung A (2019) Encouraging Participation. Discussion draft. Available at: https://ash.harvard.edu/files/ash/files/encouraging_participation.pdf

Kosack S, Creighton J, Tolmie C, Conteh F, Englin E, Gassama L, Hilligoss H, et al (2021) Brokering Collaboration: Involving Officials in Community Scorecard Programs. Ash Center for Democratic Governance and Innovation, Harvard Kennedy School. Available at: https://ash.harvard.edu/publications/brokering-collaboration-involving-officials-community-scorecard-programs

Kremer M, Gallant S, Rostapshova O and Thomas M (2019) Is Development Innovation a Good Investment? Which Innovations Scale? Evidence on social investing from USAID's Development Innovation Ventures. Available at: https://scholar.harvard.edu/files/kremer/files/sror_div_19.12.13.pdf

Larsen J (2015) Uganda: Winning Human Resources for Health. Case Study. Washington, DC: International Budget Partnership.

Lavers T and Hickey S (2015) Investigating the Political Economy of Social Protection Expansion in Africa: At the Intersection of Transnational Ideas and Domestic Politics. ESID Working Paper 47, Effective States and Inclusive Development Research Centre, University of Manchester.

Leftwich A (2009) Redefining Politics (vol. 45). London: Routledge.

Lemire S, Kwako A, Nielsen S B, Christie C A, Donaldson S I and Leeuw F L (2020) What is this thing called a mechanism? Findings from a review of realist evaluations. New Directions for Evaluation 2020(167): 73-86.

Levy B (2014) Working with the Grain: Integrating Governance and Growth in Development Strategies. New York: Oxford University Press.

Levy B and Walton M (2013) Institutions, Incentives and Service Provision: Bringing Politics Back In. ESID Working Paper 18. Manchester, UK: School of Environment and Development, University of Manchester. Available at: https://ssrn.com/abstract=2386655

Lindström, Camilla (2019) *Weak States, strong elites and acquiescent donors: State-building and aid relationships in the Democratic Republic of Congo 2006-2016.* Doctoral thesis (DPhil), University of Sussex.

Locke R M and Thelen K (1995) Apples and Oranges Revisited: Contextualized Comparisons and the Study of Comparative Labor Politics. Politics & Society 23(3): 337-367.

Lodenstein E, Dieleman M, Gerretsen B and Broerse J (2016) Health Provider Responsiveness to Social Accountability Initiatives in Low-and Middle-Income Countries: A Realist Review. Health Policy and Planning 32: 125–140.

Lodenstein E, Ingemann C, Molenaar J M, Dieleman M, Broerse J E W (2018) Informal social accountability in maternal health service delivery: A study in Northern Malawi. PLoS ONE 13(4): e0195671.

Long J and Panday P (2020) Evaluation Results of 'citizen voice & action' NOBO JATRA Program, World Vision Bangladesh USAID. World Vision. Unpublished manuscript.

Mahoney J (2000) Path dependence in historical sociology. Theory and Society 29(4): 507-548.

Mahoney J and Rueschemeyer D (Eds.) (2003) Comparative Historical Analysis in the Social Sciences. Cambridge: Cambridge University Press.

Mahoney J and Thelen K (2010) A Theory of gradual Institutional Change. In Mahoney J and Thelen K (Eds.) Explaining institutional change: Ambiguity, agency, and power. Cambridge, UK: Cambridge University Press.

Masaki T, Custer S, Eskenazi A, Stern A and Latourell R (2017) Decoding Data Use: How do leaders source data and use it to accelerate development? Williamsburg, VA: AidData at William & Mary.

Masset E and White H (2019) To Boldly Go Where No Evaluation Has Gone Before: The CEDIL Evaluation Agenda. CEDIL Paper: London.

Mayka L (2019) Building Participatory Institutions in Latin America: Reform Coalitions and Institutional Change. Cambridge: Cambridge University.

Mayne J (2015) Useful theory of change models. Canadian Journal of Program Evaluation 30(2): 119-142.

McDonnell E M (2020) Patchwork Leviathan: Pockets of Bureaucratic Effectiveness in Developing. Princeton, New Jersey: Princeton University Press.

McAdam D, Tarrow S and Tilly C (2001) Dynamics of Contention. New York: Cambridge University Press.

McAnnally-Linz H, Park B and Rajkotia R (2021) Putting Evidence to Use. Stanford Social Innovation Review Fall. Available at: https://ssir.org/articles/entry/putting_evidence_to_use#

McGee R and Gaventa J (2011) Shifting Power? Assessing the Impact of Transparency and Accountability Initiatives. IDS Working Paper, Institute of Development Studies, Brighton.

Mendiburu M (2020) La Participación Ciudadana en las Entidades Fiscalizadoras Superiores en América Latina: ¿Avance o impase? Text Prepared for the Seminario Internacional de Participación Ciudadana en Procesos de Fiscalización, Ciudad de México.

Mills L C (2019) Making the Budget Work for Ghana: Final Evaluation (English). Washington, DC: World Bank Group. Available at: http://documents.worldbank.org/curated/en/418131607356580690/Making-the-Budget-Work-for-Ghana-Final-Evaluation

Mills L (2020) All hands on deck: Harnessing accountability through external public audits. An assessment of national oversight systems. Available at: *https://www.internationalbudget.org/wp-content/uploads/ibp-idi-harnessing-accountability-november-2020.pdf*

Mohanan M, Rajan V S, Swanson K and Thirumurthy H (2020) Information and Facilitation Interventions for Accountability in Health and Nutrition: Evidence from a Randomized Trial in India. Economic Research Initiatives at Duke (ERID) Working Paper 295, February 2020.

Naeve K, Fischer-Mackey J, Puri J, Bhatia R and Yegbemey R (2017) Evaluating Advocacy: An Exploration of Evidence and Tools to Understand What Works and Why. 3ie Working Paper 29, International Initiative for Impact Evaluation (3ie), New Delhi.

Nelson E, Bloom G, and Shankland A (2018) Introduction: Accountability for Health Equity: Galvanising a Movement for Universal Health Coverage, *IDS Bulletin*, 49 (2): 1 – 18.

Nxumalo N, Gilson L, Goudge J, Tsofa B, Cleary S, Barasa E and Molyneux S (2018) Accountability mechanisms and the value of relationships: Experiences of front-line managers at subnational level in Kenya and South Africa. BMJ Global Health 3(4): e000842.

Oliver S (2019) Stakeholder engagement for navigating evidence and uncertainty. UCL EPPI-Centre, for Norwegian Ministry of Education, March 2019.

Olken B (2007) Monitoring Corruption: Evidence from a Field Experiment in Indonesia. Journal of Political Economy 115(2): 200–249.

O'Meally S C (2013) Mapping Context for Social Accountability: A Resource Paper. Social Develop-ment Department, World Bank, Washington, DC.

Ostrom E (1990) Governing the commons: The evolution of institutions for collective action. Cambridge, UK: Cambridge University Press.

Ostrom E (1996) Incentives, Rules of the Game, and Development. In Proceedings of the Annual World Bank Conference on Development Economics 1995, 207–34. Washington, DC: The World Bank.

Ostrom E (1998) A Behavioral Approach to the Rational Choice Theory of Collective Action. American Political Science Review 92(1): 1–22.

Pawson R (2013) The science of evaluation: a realist manifesto. London: Sage.

Pawson R and Tilley N (1997) Realistic Evaluation. London: Thousand Oaks and Sage.

Pierson P (2004) Politics in Time: History, Institutions and Social Analysis. Princeton, New Jersey: Princeton University Press.

Poli M, Guerzovich M F and Fokkelman E (2020) Aligning Actors for Learning: The Evolution of Social Accountability in the Dominican Republic's Education Sector. Global Partnership for Social Accountability Note. World Bank, Washington, DC. Available at: https://gpsaknowledge.org/knowledge-repository/aligning-actors-for-learning-in-the-dominican-republic-the-evolution-of-social-accountability/.

Poli M and Guerzovich M F (2020) Capacity and Implementation Support Area: Portfolio Performance Review. Global Partnership for Social Accountability Note 15. World Bank, Washington, DC. Available at: https://gpsak-nowledge.org/knowledge-repository/capacity-and-implementation-support-area-portfolio-performance-review/

Poli M, Meknassi S, Thindwa J, Kumagai S, Cavatore M and Jespersen A (2020). Collaborative Social Accountability in Fragile Settings: Lessons from Guinea, Niger, and Tajikistan. Governance Notes 24. Washington, DC: World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/34077

Poteete A R, Janssen M A and Ostrom E (2010) Working together: collective action, the commons, and multiple methods in practice. Princeton, New Jersey: Princeton University Press.

Pritchett L (2021) Let's Take the Con Out of Randomized Control Trials in Development: The Puzzles and Paradoxes of External Validity, Empirically Illustrated. Oxford Blavatnik School of Government, Oxford. Available at: https://lantpritchett.org/wp-content/uploads/2021/05/taking-the-con-out-of-rct_may2021.pdf

Przeworski A and Teune H (1970) The logic of comparative social inquiry. New York: Wiley-Interscience.

Punton M, Vogel I, Leavy J, Michaelis C and Boydell E (2020) Reality Bites: Making Realist Evaluation Useful in the Real World. CDI Practice Paper 22, March 2020, Centre for Development Impact.

Putnam R D, Nanetti R and Leonardi R (1993) Making democracy work: civic traditions in modern Italy. Princeton, New Jersey: Princeton University Press.

Raffler P, Posner D and Parkerson D (2019) The Weakness of Bottom-Up Accountability: Experimental Evidence from the Ugandan Health Sector. IPA Working Paper, Innovations for Poverty Action, Washington, DC. Available at: https://www.poverty-action.org/publication/weakness-bottom-accountability-experimental-evidence-ugandan-health-sector.

Richards R (2019) The Value of Theory of Change in Large-Scale Projects and Complex interventions. K4D Helpdesk Report, November 2019. Available at: https://assets.publishing.service.gov.uk/media/5df7bbdf40f0b60950ba8326/702_ToC_Development_for_Capacity_Building_and_Behavioural_Change.pdf

Rocha Menocal A and Sharma B (2008) Joint Evaluation of Citizen's Voice and Accountability Report: Synthesis Report. London, UK: Department for International Development.

Roche C (2015) Identifying rebels with a cause (and effect). Available at: https://www.dlprog.org/opinions/identifying-rebels-with-a-cause-and-effect

Rogers E M (1995) Diffusion of innovations (4th ed). New York: Simon & Schuster.

Samuel J, Flores W and Frisancho A (2020) Social exclusion and universal health coverage: health care rights and citizen-led accountability in Guatemala and Peru. International Journal for Equity in Health, 19(1): 1-9.

Sartori G (1970) Concept Misformation in Comparative Politics. The American Political Science Review 64(4): 1033-1053

Schaaf M, Falcao J, Feinglass E, Kitchell E, Gomes N and Freedman L (2020) We all have the same right to have health services: A Case Study of Namati's Legal Empowerment Program in Mozambique. BMC Public Health 20: 1084.

Scriven M (2008) A Summative Evaluation of RCT Methodology: An Alternative Approach to Causal Research. Journal of MultiDisciplinary Evaluation 5(9): 11-24.

Sidel J T and Faustino J (2019) Thinking and Working Politically in Development: Coalitions for Change in the Philippines. The Asia Foundation, Pasig City. Available at:

https://asiafoundation.org/wp-content/uploads/2020/07/Thinking-and-Working-Politically-in-Development_Coalitions-for-Change-in-the-Philippines_Faustino_Sidel.pdf

Squires F, Martin Hilber A, Cordero J P, Boydell V, Portela A, Lewis Sabin M (2020) Social accountability for reproductive, maternal, newborn, child and adolescent health: A review of reviews. PLoS ONE 15(10): e0238776.

Stachowiak S (2013) Pathways for Change: 10 Theories to Inform Advocacy and Policy Change Efforts. Center for Evaluation Innovation. Available at: https://www.evaluationinnovation.org/wp-content/uploads/2013/11/Pathways-for-Change.pdf

Stern E, Stame N, Mayne J, Forss K, Davies R and Befani B (2012) Broadening the Range of Designs and Methods for Impact Evaluations. DFID Working Paper 38, April 2012.

Tembo F (2012) Citizen voice and state accountability. Towards theories of change that embrace contextual dynamics. Working Paper 343, Overseas Development Institute, London.

Templeton Dunne J, Lesyna K, and Zaret A (2017) The Role of Human Rights Litigation in Improving Access to Reproductive Health Care and Achieving Reductions in Maternal Mortality, BMC Pregnancy and Childbirth 2017, Suppl 2: 367.

Tendler J (1997) Good Governance in the Tropics. Baltimore and London: The Johns Hopkins University Press.

Teskey G (2021) Accountability in PNG: Green for Go or Red for Stop? Available at: https://abtassocgovernancesoapbox.files.wordpress.com/2021/03/accountability-in-png-gt-17th-march-2021-blog.pdf

Thelen K (1999) Historical Institutionalism in Comparative Politics. Annual Review of Political Science 2: 369-404.

Thelen K (2004) How institutions evolve: The political economy of skills in Germany, Britain, the United States, and Japan. Cambridge, UK: Cambridge University Press.

Tilly C (2004) Social Movements, 1768–2004. Boulder, Colorado: Paradigm Publishers.

Tonn G (forthcoming) Development policy and impact evaluation: Contribution analysis for learning and accountability in private sector development. In Zafarullah H and Huque A S (Eds.) Handbook of Development Policy. Cheltenham, UK: Edward Elgar Publishing.

Tsai L, Morse B, Toral G and Lipovsek V (2019) Information and Accountability: Evidence Synthesis of Within Government and Citizen-Government Accountability Pathways. Washington, DC: Transparency and Accountability Initiative.

Usmani F, Jeuland M and Pattanayak S K (2021) NGOs and the Effectiveness of Interventions. Ruhr Economic Papers 902. Available at: https://en.rwi-essen.de/publikationen/ruhr-economic-papers/1137/

Vargas R (2013) Sistematización de la Iniciativa de Vigilancia Ciudadana de la Calidad de los Servicios de Salud en las provincias de Ayaviri y Azángaro – Puno. CARE International.

Waddington H, Sonnenfeld A, Finetti J, Gaarder M and Stevenson J (2019) Does Incorporating Participation and Accountability Improve Development Outcomes? Meta-Analysis and Framework Synthesis. 3ie Systematic Review 43, International Initiative for Impact Evaluation (3ie), London.

Wadeson A, Monzani B and Aston T (2020) Process Tracing as a Practical Evaluation Method: Comparative Learning from Six Evaluations. Available at: https://mande.co.uk/wp-content/uploads/2020/03/Process-Tracing-as-a-Practical-Evaluation-Method_23March-Final-1.pdf

Walker D (2018) Leveraging communities' capabilities to increase accountability for health rights: the case of citizen voice and action. Journal of Human Development and Capabilities 19(2): 181-197.

Westhorp G, Walker B, Rogers P, Overbeeke N, Ball D and Brice G (2014) Enhancing Community Accountability, Empowerment and Education Outcomes in Low And Middle-Income Countries: A Realist Review. London, UK: EPPI-Center, Social Science Research Unit, Institute of Education, University of London.

World Bank (2003) World Development Report 2004: Making Services Work for Poor People. Washington, DC: World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/5986

World Bank (2017) World Development Report 2017: Governance and the Law. Washington, DC: World Bank. Available at: https://elibrary.worldbank.org/doi/abs/10.1596/978-1-4648-0950-7.

World Bank (2018) Engaging Citizens for Better Development. Independent Evaluation Group, Washington, DC: World Bank.

White H and Phillips D (2012) Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework. 3ie Working Paper 15, International Initiative for Impact Evaluation (3ie), London.

WHO (World Health Organization) (2010) World health report 2010: Health systems financing: The path to universal coverage. Geneva: WHO.

Yosef S and Goulden J (2016) Commitments and accountability: Peru's unique nutrition journey. In Gillespie S, Hodge J, Yosef S and Pandya-Lorch R (Eds.) Nourishing millions: Stories of change in nutrition. (125-132). Washington, DC: International Food Policy Research Institute (IFPRI).

Annex 1. Common limitations of process tracing and responses to these

There are three key limitations when employing process tracing. The first limitation which Hay (2016: 500) notes, is the *misidentification* of causal mechanisms. Hay rightly points out that in order to trace processes you must first have identified them. This is a limitation for all theory-based methods which assess changes with a high degree of causal complexity. However, one of our key assumptions for the research related to the prevalence of "policy layering" provides a good point of departure for identification of outcome and key entry points for assessing processes (see Thelen and Mahoney, 2010).

A second key limitation is epistemic. Debate remains regarding whether process tracing is necessarily positivist or not. Alternative theory-based methods such as realist evaluation rightly point out a series of shortcomings with a positivist epistemology (Pawson and Tilly, 1997; Pawson, 2013). This is not a direct critique of process tracing. However, prominent process tracing theorists tend to view the method through a neo-positivist lens (Beach and Pedersen, 2019). One limitation of such a lens is that process tracing may fail to clearly identify an agent's "reasoning" and thus the motivational impulses of decision-makers (see Pawson and Tilly, 1997; Pawson, 2013). The researchers are keenly aware of this potential shortcoming, given that the origins of this research lie in a recent realist evaluation of social accountability in the health sector in Indonesia (Ball and Westhorp, 2017). A number of the research team were part of this evaluation. So, considerable efforts have been made to identify how to explicitly consider actors' reasoning as a constituent element of the approach to process tracing employed, as discussed by one of the researchers in a webinar for the Centre for Development Impact (Aston, 2021).

As Beach and Pedersen (2019) discuss, for psychological processes, we assess the mechanism in an indirect fashion through proxies (indicators) for the observable implications (Janis, 1982 in Beach and Pedersen, 2019). Hence, observable evidence provides only a partial view of the real phenomena. In this sense, process tracing refers to observable evidence of traces left by each part of the mechanism in between cause (e.g. meteor collision) and outcome (extinction). It need not refer only to observable evidence. Indeed, as recent research has shown, it is entirely possible to adopt a discursive institutionalist lens to process tracing (see Lavers and Hickey, 2015; Hickey and Hossain, 2019). The study's sensitivity to ideational anchoring and focus on policy layering should thus help diminish legitimate critique of process tracing as lacking ontological depth. However, it is worth noting that the same issues of uncovering unobserved elements which realist evaluation itself attempts to confront are equally relevant to process tracing (see Lemire *et al*., 2020).

A third key limitation relates to the potential for various biases commonly found in small-n research (White and Philips, 2012). As White and Philips (2012) note, greater effort is required to ensure that small-n research and evaluation minimize the biases which are likely to arise from the collection, analysis and reporting of qualitative data. Recent learning from applying process tracing in six evaluations by one of the researchers in the project suggests that when evaluations are partner-led or participatory, that proximity and confirmation biases can be

strong. The study found that evaluation stakeholders will tend to develop causal chains which reflect the greatest intensity of effort rather than where there is necessarily better evidence of influence or leverage (Wadeson, Monzani and Aston, 2020). Moreover, White and Philips (2012: 25) note that courtesy bias has been found to be stronger in Asian countries. So, this also bears due consideration, as three of the cases are in Asia (Indonesia and Cambodia). Employing informal Bayesian logic and assessment of evidence based on probative value, as well as facilitating the explicit assessment of the threat of rival hypotheses are the main ways in which this research will help to reduce such potential biases.

www.CEDILprogramme.org