

# **Social Accountability and Service Delivery Effectiveness: What is the Evidence for the Role of Sanctions?**

## **Background Paper**

*Tom Aston and Grazielli Faria Zimmer Santos (2022)*

Contact: [thomasmtaston@gmail.com](mailto:thomasmtaston@gmail.com)

## Table of Contents

<b>Executive Summary</b> .....	<b>2</b>
<b>Introduction</b> .....	<b>3</b>
<b>Defining accountability</b> .....	<b>5</b>
<b>Sanctions in theories of change</b> .....	<b>7</b>
<b>Methodology</b> .....	<b>9</b>
<b>Review questions</b> .....	<b>9</b>
<b>Definitions</b> .....	<b>9</b>
<b>Inclusion criteria</b> .....	<b>10</b>
<b>Methods</b> .....	<b>10</b>
<b>Study limitations</b> .....	<b>12</b>
<b>Key findings from evidence reviews</b> .....	<b>13</b>
<b>Findings on sanctions from individual studies</b> .....	<b>15</b>
<b>Use of formal and social sanctions</b> .....	<b>15</b>
Formal sanctions .....	15
Social sanctions.....	16
<b>How effective were initiatives in contributing to impact?</b> .....	<b>18</b>
<b>Sustainability and systemic effects</b> .....	<b>21</b>
<b>Negative effects</b> .....	<b>22</b>
<b>Different sanctions for different stakeholders</b> .....	<b>23</b>
<b>Collaboration and/or confrontation?</b> .....	<b>24</b>
<b>What contextual factors mattered most?</b> .....	<b>29</b>
<b>Conclusions</b> .....	<b>31</b>
<b>Recommendations</b> .....	<b>32</b>
<b>Bibliography</b> .....	<b>35</b>

**Acknowledgements:** Thanks to Jeff Thindwa, Helene Grandvoinnet, Luiza Nora, Najat Yamouri, Florencia Guerzovich, Cathy Shutt, and David Jacobstein for providing valuable comments.

## Executive Summary

Understanding how civil society can get government to respond to their needs, preferences and demands, and deliver goods and services to citizens is a central concern in social accountability initiatives. It is widely argued that sanctions make a key difference to service delivery outcomes, and that without them, transparency and accountability interventions are less likely to be effective and less likely to be sustainable (Anderson et al., 2020; Arugay, 2016; Fox, 2020; Goetz and Jenkins, 2005; Grandvoinet et al., 2015; McGee and Gaventa, 2011; Molina et al., 2017; Joshi, 2010; Joshi, 2017; Tsai *et al.* 2019). In this paper, sanctions refer to the threat or imposition of a punishment for transgressing a rule or norm.

Yet, what evidence is there to support the claim that sanctions are king? How much do we actually know about social and formal sanctions and their effectiveness in improving service delivery? Looking at 11 meta-reviews and 35 cases, this background paper sheds light on these questions and the conditions under which sanctions promoted within social accountability interventions may have contributed to improved service delivery.

Sanctions, both social and formal, feature very prominently in the scholarly literature related to accountability, so much so that some scholars have argued that it has become synonymous with punishment (Mansbridge, 2014; Schedler, 1999). While it may not be true that these concepts are fully synonymous, this trend in scholarship significantly influenced the thinking of donors such as the Department for International Development (DFID) and the World Bank over the last two decades (Grandvoinet et al., 2015; Malena and McNeil, 2010; Moore and Teskey, 2006; World Bank, 2003). As a result, sanctions also feature in the dominant models, or theories of change, in the social accountability sector. Yet, in reviewing 11 meta-reviews in the transparency and accountability sector we found that while there is plenty of theoretical argument asserting the potential effectiveness of social and formal sanctions, there was limited empirical evidence to support the claim that sanctions were key.

Our review of 35 cases revealed five mechanisms of change related to social and formal sanctions. These were: (i) “sticks” – response to punishment; (ii) “big brother is watching” – response to threat of formal punishment; (iii) diagonal accountability – response to threat of formal punishment by horizontal accountability agencies; (iv) litigation – response to legal investigations or lawsuits supported by community paralegals and legal aid organizations and; (v) response to “naming and shaming” by civil society and/or media.

We challenge the dominant view in scholarship that harder social sanctions and enforcement of stronger formal sanctions are either necessary to the achievement of higher-level service delivery outcomes or that they will deliver better and more sustainable outcomes. We find that both social and formal sanctions *can* contribute to improving service delivery outcomes across a variety of country contexts. In half of the 35 cases reviewed, we were able to establish a *likely* link between social or formal sanctions and intermediate effects over the short term with some degree of confidence. These effects ranged from increased service provider awareness and motivation, increased availability of funding, staff, and materials, to improved infrastructure quality, and in a minority of cases impact level changes such as improved test scores.

However, the role of sanctions in delivering outcomes was often unclear, outcomes were almost never sustainable, and in close to half of the 35 cases reviewed there were substantial negative effects. These effects ranged from reducing transparency and funding, to discrediting, relocating, and reprisals for advocates and whistle-blowers, threats of violence to collaborating government actors, damaging staff morale, reducing attendance, and generating conflict among staff and between staff and community members and between staff and patients, damaging trust. There are therefore some serious ethical dilemmas associated with sanctioning efforts which need to be carefully considered. We argue

that imposing sanctions without building relationships or systems to promote good behavior is unlikely to improve service delivery outcomes in a sustainable way.

Another, perhaps surprising, finding was that there are some actors that are regular targets of sanctions, and in many cases, these actors are a lot weaker than is commonly assumed. Three broad types of actors were the most common targets of punishment: (i) absentee nurses or teachers who had their pay or allowances reduced; (ii) offending officials who were either suspended, relocated, or fired; and (iii) contractors who had to cover the cost of rejected materials or faced lawsuits, alongside civil servants involved in contracting.

We suggest that closer relationships may perhaps reduce stakeholders' appetite to impose sanctions. Particularly in the health sector, we found that more proximate relationships created disincentives for confrontation, and in such circumstances, a "policing" approach to monitoring was also deemed inappropriate and counterproductive. Conversely, it seems that actors generally prefer to sanction "others," i.e., when an actor/organization was outside the group. Short-term consultants, contractors, and suppliers were easy (and quite vulnerable) targets for harder sanctions.

It has also been argued that there may be productive combinations of collaborative and confrontational tactics—i.e., hybrids. We found that many supposed confrontational and collaborative hybrids were, in fact, dislocated in time and space. Many so-called "inside-outside" strategies, therefore, seem to be a potential mischaracterization. We also found, as Fung and Wright (2003) argued nearly two decades ago, that adversarial forms of engagement cannot easily be redeployed for collaborative purposes.

It is widely argued that supposedly 'weaker' forms of citizen engagement are less effective than those with 'strong enforceability (McGee and Gaventa, 2011).' We found no compelling evidence to support this contention. A quarter of cases reviewed were collaborative for certain periods or in certain locations. On average, these were slightly more successful when compared with periods or locations when imposing sanctions was a key strategic emphasis. So, soft power can also be powerful. Rather than one approach necessarily being superior (confrontational, collaborative, or hybrid) however, we argue that the best approach is likely to be the one most appropriate to the context.

While we were unable to identify strong trends of contextual factors which enabled social and formal sanctions to play a role in enhancing service delivery, we were able to identify several conditions which we believe offer the greatest promise. These conditions were: (i) supportive leaders who played a role as champions; opening doors or accompanying civil society efforts; (ii) high capacity and legal authority of oversight agencies; (iii) competitive elections, which provided windows of opportunity for CSOs to combine political and social accountability efforts, and; (iv) vulnerable public servants and service providers already in relatively precarious situations and are thus easy targets.

Overall, our study finds that sometimes sanctions can be effective, but punishment is not the answer to all the world's problems. Given these limitations, we recommend that it is time to reconsider "carrots" and enquire further into the enabling conditions for bureaucrats and service providers. Relatedly, scholars, evaluators, and program teams should look more closely at service providers' or civil servants' motivations and take context into account more seriously. To uncover these contextual and motivational features, we believe that scholars, evaluators, and programeams should also make better use of theory-based and participatory methods. And perhaps most importantly, donors and practitioners should carefully consider and mitigate the potential for backlash from sanctions-based approaches.

## **Introduction**

Reviews and syntheses of the effectiveness of transparency and accountability interventions suggests that the evidence is mixed (ePact, 2016; Gaventa and Barrett, 2010; Fox 2014; Lodenstein et al., 2016; McGee and Gaventa 2011; Kosack and Fung, 2014; Rocha Menocal and Sharma, 2008; Tsai *et al.* 2019;

Waddington et al., 2019; Westhorp *et al.* 2014). This has promoted scholars to question why some interventions showed positive results in improved service delivery (Alhassan et al., 2016; Björkman and Svensson, 2009; Gullo *et al.* 2017; Mohanan *et al.* 2020) whereas others did not (Arkedis *et al.* 2019; Raffler et al., 2019). To date, no fully satisfactory answer to this question has been forthcoming.

In the past decade, a growing set of stakeholders have argued that mixed outcomes are, in large part, the result of weak or inexistent sanctions. For example, a recent evidence synthesis from Tsai *et al.* (2019) argues that success is more likely in contexts with relatively stronger top-down sanctioning of local officials by those in higher levels of government. Many prominent scholars have advocated for greater international incentives and support for CSOs' capacity to sanction or capacity to trigger state sanctions. Capacity to sanction most commonly refers to governmental capacity to apply negative sanctions (legal or administrative), but it also often refers to pressure from below generated by different efforts to litigate, to name and shame service providers and public authorities or protest – i.e., social sanctions (Anderson et al., 2020; Fox, 2014, 2016, 2020; Joshi, 2010; McGee and Gaventa, 2011).

However, scholarship on the effectiveness of sanctions on service delivery outcomes offers a contradictory picture. Westhorp et al.'s (2014: 69, 70) realist synthesis of accountability in the education sector found 'few examples where sanctions had actually been applied' and 'very limited evidence of the application of such sanctions,' thus raising questions about whether the aforementioned hopes are based on evidence, or faith. Another systematic review of the field which looked across sectors (except education), questioned the effectiveness of confrontational approaches and challenged the contention that social sanctions are likely to be effective in changing the balance of power between service providers and service users (Waddington et al., 2019).

It thus seems reasonable to ask what evidence is there that sanctions are either a necessary condition for success or a key contributory factor towards achieving success. It is commonly recommended to employ a combination of tactics at multiple levels, bridging supply and demand (Fox, 2014; Gaventa and Barrett, 2010; McGee and Gaventa, 2011; Joshi, 2017; Rocha Menocal and Sharma, 2008). However, relatively little attention has been paid to demonstrating *how* these tactics interact, and under what conditions for instance, collaborative and confrontational tactics are complementary or contradictory. A focus on sanctions enables us to unwrap intervention packages and draw attention to the need for scholars to demonstrate these connections empirically and show their contribution to service delivery outcomes.

This paper intends to shed light on the role of social and formal sanctions in the social accountability sector. It will reveal how prominently they feature in theories of change and evidence reviews, and it will appraise the available evidence on their effectiveness in improving service delivery outcomes. We also identify several mechanisms through which, under some conditions, sanctions play an important role and can contribute to these outcomes. We hope that this effort will enable scholars and practitioners to more closely scrutinize when, how, and why sanctions may make a difference (or not) in particular contexts.

The structure of the paper is as follows. We will first situate the review in a theoretical context and explain how sanctions rose to prominence in the last few decades. We will then consider how a focus on sanctions has influenced theories of change and the assumptions which underpin the strategies deemed to be most effective in achieving service delivery outcomes. Following a brief explanation of the review's methodology, we will look at the evidence base on 11 meta-reviews of the accountability sector. We will then consider cross-sectoral trends related to sanctions and different mechanisms of change linked to social and formal sanctions across 35 case studies, as well as the positive and negative effects of sanction-based approaches. We will look at how different sanctions are tailored to different stakeholders. We discuss the potential of hybrid approaches that combine collaborative and confrontational approaches and whether harder forms of accountability are more effective or not. We then explore potentially important contextual factors to support sanction-based approaches, and we end the paper by providing a conclusion with some implications and recommendations for practitioners.

## Defining accountability

There is no consensus in either scholarship or practice on what “accountability” means, and thus a variety of definitions of accountability exist (Bovens et al., 2014; Honig and Pritchett, 2019). Yet, perhaps the most cited definition as theoretical justification in transparency and accountability scholarship comes from Andreas Schedler’s (1999: 14 – 15) definition of *political* accountability. Schedler sees accountability as a combination of answerability and enforcement. Answerability is about citizens asking uncomfortable questions and calling into question those in authority who are compelled to provide answers to those questions, to inform and explain their actions. However, for Schedler (1999: 14 – 15), answerability is insufficient:

‘Accounting actors [should] not simply “call into question” but also eventually *punish* improper behavior and accordingly, that accountable persons not only tell what they have done and why, but bear the consequences for it, including eventual negative sanctions. Exercises of accountability that expose misdeeds but do not impose material consequences will usually appear as weak, toothless, “diminished” forms of accountability. They will be regarded as acts of window dressing rather than real restraints on power (our emphasis).’

In Schedler’s conceptualization, misdeeds have already been exposed, and it is assumed that an abuse of power has already occurred. There is also an assumed sequence from questioning to punishing in this definition. Accountability is thus defined by its assumed absence. Schedler’s (1999: 15 – 16) emphasis on ‘punishment,’ ‘sanctions,’ ‘impeachment,’ ‘destruction of reputation through public exposure,’ ‘dismissal’ or ‘legal sanctions’ set a particular tone for accountability. ‘Weak,’ ‘toothless,’ and ‘diminished’ are also not merely factual statements related to a lack of force, but also *normative* statements describing an undesirable state. Alongside providing information and justifying decisions, sanctions are widely considered a constitutive element of accountability in scholarship (Behn, 2001; Bovens, 2005; Cunil-Grau, 2008; Fox, 2007, Fox, 2014; Goetz and Jenkins, 2005; Herrera and Mayka, 2019; Muligan, 2000; Peruzzotti and Smulovitz, 2006; Strom, 2000, 2003). Jane Mansbridge (2014: 56) notes the historical trend over the last several decades from a definition of accountability in French (*rendre compte*), German (*Rechenschaft abgeben*), and English (give account of)<sup>1</sup> which is broadly synonymous with answerability to one which has slowly begun to ‘connote sanctions (Mansbridge, 2015: 56).’ Mansbridge even argues that ‘accountability has become synonymous with punishment (Mansbridge, 2014: 55).’ While this may be a slight exaggeration, it nonetheless suggests there has been a potential overcorrection.

This discursive shift is significant for *social* accountability because Schedler’s emphasis on political accountability, its primary actors, and concomitant assumptions about their motivation and behavior has been adopted in some definitions of social accountability. For example, Enrique Peruzzotti and Catalina Smulovitz (2006: 10) define social accountability as a ‘nonelectoral yet vertical mechanism of control of *political* authorities (our emphasis).’ They argue that:

‘Exposing and denouncing cases of government wrongdoing... generates symbolic costs to the officials or agencies suspected of wrongdoing but also, by bringing cases of corruption or official misconduct into the public agenda, forces political institutions to address these cases and raises the actual costs of illegal or improper political behavior (Peruzzotti and Smulovitz, 2006: 11).’

---

<sup>1</sup> See also Fox (2018) for a discussion of *control social* (social control) and *rendición de cuentas* as a translation of the term “accountability” in Spanish.

As in Schedler, then, the reader is predisposed to assume government wrongdoing has already occurred and that corrupt politicians (as primary targets) and their officials must therefore be punished for their wrongdoing.<sup>2</sup>

The lion's share of scholarship and work from major donors such as DFID and the World Bank effectively adopted Schedler's definition, its sequencing, and its hierarchy (see Grandvoinet et al., 2015; Malena and McNeil 2010; Moore and Teskey, 2006; Staplehurst and O'Brien, 2005; World Bank, 2003). Some scholars retained scope for rewarding good behavior (carrots) and supportive relationships of civil servants (Ackerman, 2005; Fox, 2014; Paul, 1992), however, as the initial faith in the effectiveness of transparency interventions declined in the last decade, scholars accentuated Schedler's hierarchy and its underlying assumptions, with punishment (sticks) as the primary goal.

Mick Moore and Graham Teskey (2006: 2 – 4) argued that accountability has four sequential stages: 1) standard setting to establish expected behavior; 2) investigation regarding whether standards have been met; 3) answerability to defend actions and face challenging questions, and 4) sanctions to punish actors for falling below standards. Jonathan Fox (2007) also delineated soft accountability (answerability) and hard accountability (sanctions) and argued that accountability is weak and incomplete without sanctions, compensation, and remediation (Fox, 2007, see also Muligan, 2000). Rosemary McGee and John Gaventa (2011: 11) have further critiqued the 'weaker category' of 'answerability' and 'soft peer or reputational pressure [vs] strong enforceability (see also Goetz and Jenkins, 2005).' This sense of hierarchy led Aries Arugay (2016: 8, 26), for instance, to assert that not only were sanctions a '*sine qua non* of accountability,' they were 'accountability mechanisms *par excellence*.'

One further aspect of Schedler's (1999) definition that has gained additional traction among transparency and accountability scholars in recent years, particularly in a context of shrinking civic space, is his *a priori* assumption of improper behavior. Fox (2014: 32) has argued that, as resistance is expected by the state, 'conflict should be both expected and necessary.' Thus, while "teeth" remains a far broader concept than simply sanctions, demobilizing anti-accountability forces became a centripetal component of the concept. Likewise, Anu Joshi (2017: 161) suggests that 'public accountability failures are not accidental—they occur due to embedded power structures and political dynamics that are systemically anti-accountability.' This assumed widespread impropriety and impunity (rather than of a minority of "bad apples") is thus provided as a justification for predominantly *adversarial* forms of "countervailing power" to reduce and neutralize the power advantages of ordinarily powerful actors, typically emphasizing differences between groups rather than commonalities (Fung and Wright, 2003: 260). Such an approach is increasingly associated with advocacy efforts to forge vertical and diagonal links to higher-level bodies within the accountability ecosystem – i.e., Supreme Audit Institutions and courts – which are hypothesized to provide a degree of pressure or threat of sanction that will motivate lower-level solutions (Anderson et al., 2020; Fox, 2020; Herrera and Mayka, 2019; Joshi, 2017).

Schedler's (1999) hierarchy remains largely unchallenged because "hard accountability" is seen by various scholars to contribute to social (or retributive) justice outcomes and the rebalancing of power.<sup>3</sup> In other words, sanctions are *right* in their own right. However, as will be discussed in the next section, a significant portion of scholarship argues that with a stronger emphasis on both formal sanctions and social sanctions, better *service delivery outcomes* will also be achieved.

---

<sup>2</sup> Goetz and Jenkins (2005: 9) also adopt Schedler's definition, with a more moderate view that 'answerability and enforcement are equally important' where 'both are necessary,' but 'neither is sufficient.' Yet, they argue that non-confrontational approaches to citizen engagement secure cooperation at the expense of the enforcement dimension of accountability. They argue that 'confrontation is inevitable (2005: 96),' yet this is largely due to their emphasis on accountability *failures* and *a priori* assumptions of abuses of power which underpin this, and therefore, in practice, define accountability by its assumed absence.

<sup>3</sup> See, for example, Hevia and Vergara-Lope's (2019) advocacy for a rights paradigm with a more expansive focus on sanctions in the education sector. See also Tsai (2021) for a discussion on hard accountability and retributive justice.

Notwithstanding, pursuing hard accountability is not always desirable and it may not necessarily lead to better outcomes. Mansbridge (2015: 55) notes that ‘*sanction-based* accountability is most appropriate in contexts of justified distrust. Yet, it also creates distrust’ through what Behn (2001: 83) refers to as ‘institutionalized suspicion.’ Mansbridge’s (2014) argument and that of Honig and Pritchett (2019) for “trust-based” or “deliberative” forms of accountability have yet to gain much wider traction in scholarship. Yet, their concerns regarding an appropriate balance and proportionality appear to reflect concerns from the education and health sectors, which are the most prominent sectors for social accountability interventions.

As Zeitlin et al., (2011: 21) suggest, where human resourcing is a challenge, pressing to fire even underperforming staff may not be desirable in the absence of an alternative. And as Nxumalo et al., (2018: 5) have argued, the ‘laudable intentions’ of bureaucratic accountability may have ‘unintended consequences,’ distracting staff attention, diminishing staff motivation, and even weakening trust between providers and the community. Similarly, the United Nations Educational, Scientific and Cultural Organization – UNESCO (2017) also questions the effectiveness of sanctioning schools or teachers as a basis to improve education outcomes and notes that these may even undermine overall quality and learning. Yet, as we shall see, these concerns are rarely acknowledged in theories of change in the social accountability sector.

## Sanctions in theories of change

Another important way to consider the potential role of sanctions in social accountability programming is through the lens of theories of change. Theories of change are essentially hypotheses which explain how and why we expect change to happen (e.g., if x and y are done, then we anticipate z results, because...). Sue Funnell and Patricia Rogers’ (2011: 357) seminal work on theories of change outlines what they call “archetype” theories of change (or meta-theory). Each of these meta-theories features prominently in the social accountability field: information, community capacity building, case management, direct service delivery, and carrots and sticks. To date, the field has been largely divided by two of these: information and carrots and sticks. The first of these is referred to as “information plus” interventions, as the provision of information to civil society groups forms the backbone of interventions, and training, coordination, or incentives are added to information (Tsai *et al.* 2019).

Incentives are most commonly associated with the other key archetype – “carrots and sticks.” In this archetype, behavior is influenced through incentives (carrots) to promote positive behaviors or use penalties and threats (sticks) to deter undesirable behaviors. However, Funnell and Rogers (2011) underline that carrots and sticks are best conceived as *separate* theories of change. They argue this because, in their view, ‘rewards may be given to everyone who behaves in the desired manner, whereas penalties are applied to those who are exceptions to the rule.’ Mansbridge (2014: 59) refers to this as a ‘*selection core* and a *sanction periphery*.’ For a “sticks” theory of change, in particular, Funnell and Rogers (2011: 363 – 365) note that it is important to consider the social legitimacy of regulations, an understanding of community norms, effective communication, and interrogating the various phases of the deterrence process (detection, prosecution, and application of sanctions). And yet, as we will see, few of these are seriously taken into account in the interventions (or sets of interventions) we review in this paper.

Context matters for which combination of these (or other) archetypes may work. Stephen Kossack and Archon Fung (2014: 76 – 77) developed a rubric of five ideal-typical “worlds” of transparency which speaks to the importance of context in shaping how information can help improve accountability and service delivery. These “worlds” are the following: (i) competitive services where, with good information, service users can simply choose between providers (exit options); (ii) where there may be little competition but where there are service providers are willing (*ex ante*), providing the opportunity to collaboratively develop more effective delivery practices; (iii) where there may be little competition and where providers are unwilling, so it is necessary to find ways to shift the



balance of power between citizens and providers (where social sanctions might be most effective); (iv) where there is little competition and service providers are willing but policy makers and/or politicians are willing to engage (where formal sanctions might be most effective), and (v) and where everyone is unwilling (and where protest may be the only option).

Despite Kossack and Fung (2014) suggestion of potentially different recipes for different contexts, a carrots and sticks archetype has been employed almost uniformly in the transparency and accountability field, given that theories influenced by the information archetype have increasingly focused on sanctions (known as “transparency plus”) – i.e., transparency *plus* a mixture of monitoring and social and formal sanctions (see Tsai et al., 2019: 7). Where hypothesized causal pathways are made explicit, it is this logic which predominates, and sanctions feature far more visibly than rewards (Fox, 2014; Molina et al., 2017; Tsai et al., 2019; Waddington et al., 2019). Schedler’s general perspective has therefore stuck. For Grandvoinet et al., (2015: 59), social accountability interventions seek to shift incentives either by increasing ‘potential reputational costs or the fears of triggering traditional accountability mechanisms and subsequent sanctions.’ They argue that this is the ‘main route’ to achieving outcomes such as improved public services. Anu Joshi (2010) likewise hypothesizes that:

‘Accountability or transparency mechanisms that have the potential to trigger strong sanctions are more likely to be used and be effective in improving responsiveness [...]. Without the threat of effective sanctions (and resulting impacts), citizen mobilization is difficult to sustain in the long run. Social accountability mechanisms have impact when they can trigger traditional accountability mechanisms such as investigations, inspections and audits (Joshi, 2010: 12).’

Joshi suggests that without the threat of strong, formal sanctions, transparency and accountability mechanisms are less likely to be effective and less likely to be sustainable. The assumption here is that social accountability mechanisms will not have an impact *without* strong sanctions. However, Joshi acknowledges that ‘this chain of causation is seldom explicitly examined (Joshi, 2010: 6).’ This was later paired with the hypotheses that ‘confrontational tactics on the outside [i.e., social sanctions] might shift the terms of the debate, enabling broader collaboration within,’ and that the ‘threat of confrontation nationally may [...] provide incentives for the government to collaborate (Joshi, 2017: 168).’ Both these hypotheses entail some combination of social and formal sanctions. Yet, to date, as Anderson et al., (2020) note in their recent review of DFID programming, these hypotheses remain largely untested, and thus unconfirmed. In Tsai et al’s (2019) review, of the seven steps presented in a causal mechanism, sanctions feature explicitly in three. These are related to citizens’ ability to hold government to account (knowledge of how to monitor and sanction), their motivation to hold government to account (willingness to monitor and sanction), and effort to hold government to account (effort to monitor, sanction or reward).

In Waddington et al’s (2019) systematic review, the authors highlight the appropriateness of the strength of the social sanctions expressed as a (weak) facilitating factor for merit goods such as health services or food subsidies where citizens interact directly with providers. However, in their view, these provide insufficient incentives for public goods (e.g., roads) where the transaction intensity between users and providers is less. Alternatively, in Molina et al’s (2017) systematic review, informal and formal sanctions are presented in a model as the *only* motivational impulse.

These reviews therefore suggest that under certain conditions, it is plausible that social and/or formal sanctions may provide motivational incentives for service providers and civil servants to change their behavior. This literature also asserts that harder forms of accountability (sanctions, compensation, and/or remediation) should be more effective than softer forms of accountability in improving public services. However, as David Jacobstein suggests, what are presented as causal statements are often instead simply the expression of normative assumptions which are not necessarily assessed

empirically (Jacobstein in Aston, 2020). In the following section, we detail our methodology to help get closer to unpacking these assumptions and assessing the evidence.

## Methodology

### *Review questions*

Our proposed research questions were the following:

What do meta-reviews in the transparency and accountability sector tell us about the contribution that social and formal sanctions make to service delivery/performance?

Under what conditions and how have social and/or formal sanctions contributed to enhanced service delivery/performance from social accountability initiatives?

### *Definitions*

Definitions of social accountability are not uniform (Gaventa and McGee, 2013: 9–10; Joshi and Houtzager, 2012: 150–151), and this hints at potential tensions between theory and practice. While Schedler's (1999) definition underpins many theories of change in the sector, a more common definition used by social accountability practitioners is 'an approach to accountability that relies on *civic engagement*, i.e., in which it is ordinary citizens and/or civil society organizations that participate directly or indirectly in exacting accountability (Malena et al., 2004: 2).' In this definition, interventions may include participatory budgeting, public expenditure tracking, monitoring of public service delivery (e.g., social audits and scorecards), investigative journalism, public commissions, and citizen advisory boards, among others. Such interventions often take place in concert rather than individually.

These *citizen-driven* accountability measures are supposed to complement *conventional accountability mechanisms* such as political checks and balances, accounting and auditing systems, administrative rules, and legal procedures. While we recognize multiple directions of accountability (Halloran, 2015), as per the World Development Report 2004 (World Bank, 2003), it is important to note that in this conceptualization the primary targets of citizen-driven social accountability are service providers and civil servants rather than politicians, who are instead *indirect* or secondary targets.<sup>4</sup>

It is also important to define what we mean by "increased service performance." Here we follow Tsai et al. (2019) in determining service performance as the outcome of interest. We view this as a co-terminus with "service delivery." We consider intermediate (or process) effects such as increasing the provision of information by service providers, the provision of infrastructure or materials, increases in budgets, teaching quality or quality of care, and perceptions of trust, as well as deeper human development outcomes (or impacts) such as student enrolment, attendance, retention, and learning outcomes, stunting rates, and mortality rates.

Social sanctions and formal sanctions are rarely defined or disaggregated in the social accountability literature. In English, "sanction" is an ambiguous term. It is both positive and negative; it can refer to a penalty (i.e., punishment) for disobeying a law, rule, norm, official permission or approval for an action. In this paper, we refer to negative sanctions only, as this is the most common usage in the sector.

---

<sup>4</sup> As Peruzzotti and Smulovitz's (2006) rival definition sets politicians as the primary targets of accountability, omits service providers from consideration, and does not regard improved service provision or performance as a goal, we argue that this is a form of long-route accountability – i.e., non-electoral political accountability (see Aston, 2021 for further discussion).

Social sanctions are rarely well defined in the social accountability literature. However, these will be understood as the credible threat of a penalty or punishment or the imposition of such a penalty or punishment by members of civil society. This will include efforts by citizens to “name and shame” service providers or civil servants, with the aim to damage their reputation (Rocha Menocal and Sharma, 2008: 5 – 6). This may also comprise marches, protests, riots, or other expressions of voice such as media exposés which share this aim, as well as different types of litigation pursued by citizens’ groups. We will also include in-group sanctions by reference groups who may threaten social shunning or ostracism. Formal sanctions relate to what Malena et al., (2004) term “conventional accountability mechanisms.” This form of sanctions is threatened or imposed by state actors such as the courts, auditors’ offices, ombudsmen (i.e., legal/judicial sanctions), as well as punishments imposed within government ministries or service delivery organizations (i.e., organizational, or administrative sanctions) where these may be triggered by citizen engagement efforts. So, these may include indictment, criminal conviction, firing staff, issuing fines, pay deductions, or other forms of disciplinary measures.

In terms of defining “credible” social sanctions, some degree of value judgement is inevitable. “Pressure” is a very commonly used term in both scholarship and practice, but this too escapes definition or sustained attempt at measurement. There is a supposed association documented in scholarship between higher dosage interventions – e.g., more organizations, more levels, longer duration – and higher-level effects (Fox, 2014; Joshi, 2010).<sup>5</sup> However, it would be reasonable to infer from scholarship that there ought to be some connection between the degree of pressure and the level or sustainability of outcomes achieved. Following this line of thinking, the intensity of threats of sanctions (social or formal) or severity of sanctions that are actually imposed ought to have some bearing on these outcomes. Likewise, if sanctions are to be credible, then the precedence of sanctions – i.e., the past imposition of punishment – should be an important proxy which ought to incentivize actors to change their behavior (Olken, 2007; Tsai, 2007).

### *Inclusion criteria*

The inclusion criteria for cases were the following. Firstly, the threat of sanctions or the imposition of sanctions must be able to demonstrate a direct link to civic engagement efforts. This engagement could be through forums in which government invites civil society to participate or spaces created by civil society to engage government. Secondly, as social accountability is primarily concerned with the provision of service delivery, accountability efforts which could not demonstrate a connection to specific services were excluded. Thirdly, it is important that sanctions be linked to some form of project, program, or initiative to be considered an intervention. Various efforts of social movements were included, but only if they could demonstrate such a connection. Fourthly, this study is concerned with only low- and middle-income countries, so higher-income countries were excluded. And finally, while no methodological preference guided study inclusion, cases within studies that did not meet a minimum threshold of evidence for a plausible explanation were also excluded.

### *Methods*

As McGee and Gaventa (2011) note, methodological paradigm battles may obscure various aspects of accountability work. This work unfolds in complex contexts of socio-political relations. Processes of change may be non-linear, and mechanisms of change are often difficult to identify. Theories of change are also generally implicit or invisible (see also Joshi, 2010). To address this challenge, this review employs a realist approach which prompted us to emphasize context, mechanisms, and outcomes (Pawson and Tilly, 1997; Westhorp *et al.* 2014). Rather than proposing a grand theory, a

---

<sup>5</sup> Specific dosage criteria remain a matter of debate. Tom Wein (2020) defines criteria of duration of an intervention, length and frequency of contact, range of stakeholders, breadth of accountability channels, degree of embeddedness, level of experience of staff and scale of the intervention. However, in comparing Björkman and Svensson’s (2009) intervention which detected effects with that of Raffler et al., (2019) which did not, it is unclear precisely what difference dosage makes, at least with these criteria.

realist approach aims to build middle-range theories which are specific enough to apply to particular programs but abstract enough to apply across programs. So, findings may transfer to comparable contexts but are not intended to be fully generalizable.

In terms of the review process, we adapted Hagen-Zanker and Mallett's (2013) guidance for conducting an evidence-focused literature review, which recommends two layers of review. For the first layer, the kind of studies included were transparency and accountability reviews. We reviewed the eleven most cited and largest syntheses, meta-reviews, and systematic reviews since the seminal World Development Report in 2004 (e-Pact, 2016; Fox, 2014; Gaventa and Barrett, 2010; Hanna et al., 2011; Kosec and Wantchekon, 2020; McGee and Gaventa, 2011; McNeil and Malena, 2010; Molina et al., 2017; Rocha Menocal and Sharma, 2008; Tsai et al., 2019; Waddington et al., 2019). We first considered the frequency with which keywords appeared in these eleven reviews, excluding mentions that were merely titular or referential. We compared the frequency of words approximating the semantic field for sanctions and derivations of these words alongside words which more closely approximate "answerability" to understand the relative balance of scrutiny within these studies.

For the second layer, we reviewed the studies cited in these reviews which explicitly included empirical data on sanctions. We further conducted a database search to complement the evidence uncovered in the prioritized meta-reviews. A time limit for retrieval of studies was set at a year. Search strings included the relevant keywords and synonyms related to social accountability interventions and their outcomes. Synonyms for the keywords or different kinds of interventions used, need to be included in the search string. The aim was to build upon Tsai et al's (2019) semantic search terms to maximize potential comparability.

This was complemented by snowballing strategy as well as a grey literature search. We first identified 19 experts<sup>6</sup> from both academia and practice and deliberately ensured overlap with the meta-reviews included. We contacted these experts and asked them to suggest five key publications on that question and reviewed their own publications and websites. For grey literature capture, we relied upon material located outside the orthodox peer review channels, especially through the network of the Global Partnership for Social Accountability (GPSA) and the Transparency and Accountability Initiative (TAI). There were no specific limits on the number of studies to be reviewed, given that we anticipated retrieval to be limited. In total, beyond the 11 meta-reviews, 35 cases met the inclusion criteria.

In terms of screening, we assessed the relevance of the literature included in an Excel document. This also included an evidence assessment which comprised the classification of studies included by research design and method, year of publication, geographic coverage, type of intervention, and evidence of outcome, among others. We analyzed subsets of studies, by dividing out those which referred to social sanctions or formal sanctions or where there was a combination of these. Given the limited range of contexts included in many reviews (see Tsai et al., 2019), we further analyzed whether there was variation by geographical area. Our review was able to include an appraisal of cases across 17 countries.<sup>7</sup>

With regards to assessing the quality of studies, we developed an evidence rubric based on Department for International Development's (2014) guidance on assessing the strength of evidence. Given that we were chiefly concerned with the strength of evidence itself rather than the quality of studies overall, we also saw fit to assess evidence based on its "probative value" for a mechanistic explanation. In this respect, we also looked at "evidentiary fragments" within studies which can also aid causal leverage (see Beach and Pederson, 2019; Pawson, 2007). In this study, we take a realist

---

<sup>6</sup> Thanks to Jonathan Fox, Angela Baily, Suchi Pande, Lily Tsai, Alisa Zomer, Cathy Shutt, Alina Rocha Menocal, Hugh Waddington, Jennifer Stevenson, Ada Sonnenfeld, Florencia Guertzovich, Brendan Halloran, Courtney Tolmie, Sue Cant, Gilbert Muyumbu, Steadman Noble, Rebecca Haines, Annette Fisher, and Marta Schaaf.

<sup>7</sup> Bangladesh, Colombia, Dominican Republic, Ethiopia, Ghana, Guatemala, India, Indonesia, Kenya, Macedonia, Mozambique, Nepal, Niger, Peru, Philippines, South Africa, and Uganda

definition of mechanisms as consisting of actors' reasoning in response to opportunities and resources (Pawson and Tilly, 1997). This relies on describing the:

- a) Actor(s) whose decision-making has been changed;
- b) "Reasoning" that underlies the changed decision, and;
- c) Outcome of the different decision (Westthorp et al., 2014: 39).

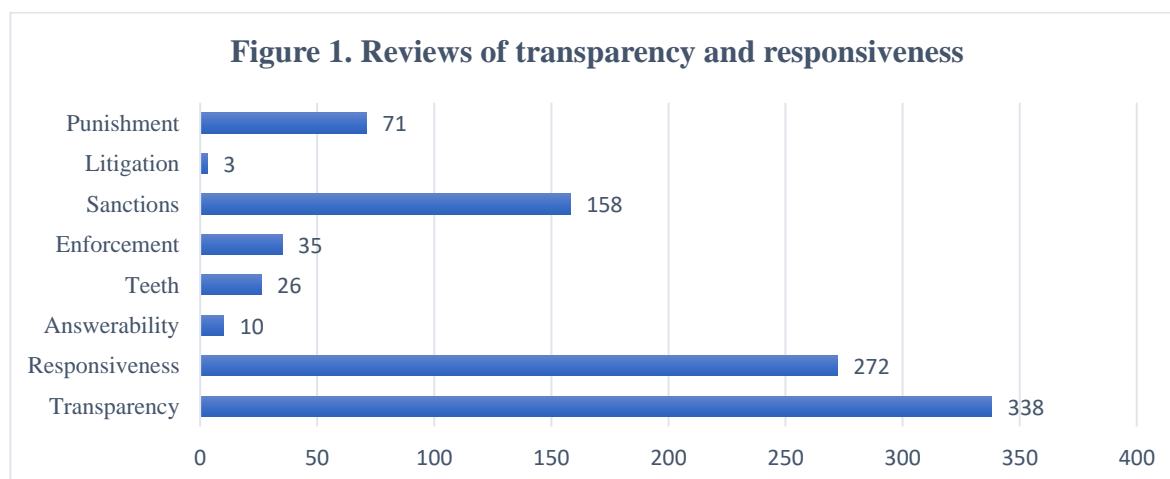
From a realist standpoint, mechanisms may act in conjunction with one another to achieve an outcome. So, a single case may have more than a single mechanism of change, and there may be interactions between these to contribute to different outcomes. Cases were iteratively coded against a set of hypothesized causal mechanisms, teasing out whether there was evidence of an intention to trigger a response via one of the proposed mechanisms, whether there was evidence of such a response, and whether there was evidence of stakeholders' reasoning which underpinned that change in behavior. The mechanisms identified in the study can be found in figure 3.

### *Study limitations*

There are two main limitations of the study. The first is the breadth of the definition of social accountability. On one hand, the study broadened the definition to expand the possibility of finding interaction effects between different directions of accountability (i.e., the potential relationship between social and formal sanctions). The exclusion of *purely* non-institutionalized forms of protest and the requirement for a deliberate link to interventions is considered justified, however, this meant that various forms of social sanction threat or action were excluded. The second related limitation is data gaps in documenting the evidence of either form of sanction threat (social and formal). We were unable to discern the degree to which this gap may be methodological (i.e., it is difficult to empirically substantiate a threat and resultant deterrence, or decouple the effects of monitoring and sanctions), evidentiary (i.e., sanction threats are applied but simply not documented), or causal (i.e., sanctions were applied but were ineffective). Scholarship acknowledges that there is a dearth of evidence on the connection between sanctions and downstream impacts. Few examples are documented where sanctions have been applied and there is limited evidence of the application of sanctions (see Tsai et al., 2019; Westthorp et al., 2014). It seems likely that this relative absence is due to a mixture of the above reasons.

## Key findings from evidence reviews

Evidence reviews since the World Development Report in 2004 have been focused on transparency and responsiveness rather than answerability or enforcement.<sup>8</sup> However, as shown below, there were a significant number of references to sanctions:



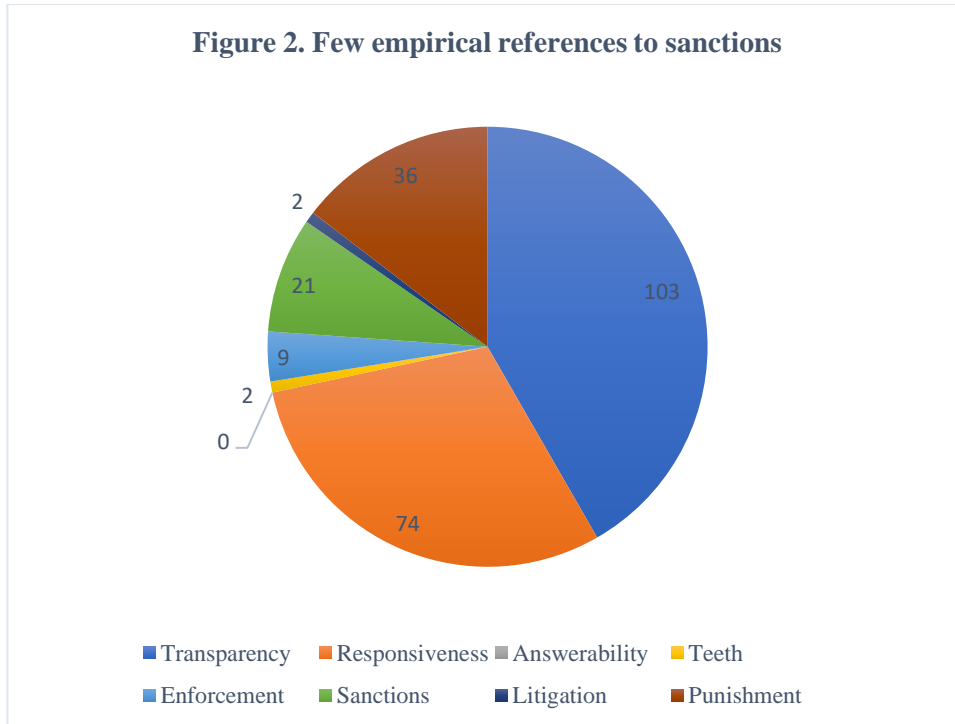
Behind these data, we also see significant variation in the emphasis of inquiry. Tsai et al., (2019) mentioned sanctions 86 times, accounting for over half of all mentions and four times more than any other review. Yet, half of the studies mentioned sanctions only a handful of times. “Teeth” is mentioned almost exclusively by Fox’s evidence review (23 mentions), and nearly all references to punishment come from Hanna et al’s (2011) anti-corruption review (62 references). The word “sanction” does not even appear in DFID’s macro-evaluation Qualitative Comparative Analysis (QCA) of 50 social accountability initiatives, nor do the words “punishment,” “enforcement,” “litigation” or “teeth.”

We found that most references to sanctions in evidence reviews were not empirical, but theoretical.<sup>9</sup> We found just 21 empirical references to “sanctions (fewer than 2 per study),” 36 references to “punishment,” only 9 references to “enforcement,” 2 references to “litigation,” and only 2 references to “teeth.” There were, however, dozens of studies which assessed transparency and responsiveness. The number of references is illustrated below:

<sup>8</sup> e-Pact, 2016; Fox, 2014; Gaventa and Barrett, 2010; Hanna et al., 2011; Kosec and Wantchekon, 2020; McGee and Gaventa, 2011; McNeil and Malena, 2010; Molina et al., 2017; Rocha Menocal and Sharma, 2008; Tsai et al., 2019; Waddington et al., 2019.

<sup>9</sup> Our focus was on multi-sectoral reviews; however, this trend also holds for other similar sector-focused evidence reviews such as in education, water, and health (Carr-Hill et al., 2016; Hepworth et al., 2020; Squires et al., 2020; Westhorp et al., 2014).

**Figure 2. Few empirical references to sanctions**



Molina et al’s (2017) systematic review provided no specific empirical references which demonstrate that either social or formal sanctions mattered, as proposed in their hypothesized causal mechanism. Tsai et al., (2019) argue that success appears more likely in contexts with relatively stronger top-down sanctioning of local officials by those in higher levels of government. However, they acknowledge that a ‘dearth of evidence makes it difficult to know whether... development outcomes are driven by greater monitoring, sanctioning, or some combination of both (Tsai et al., 2019: 25, 4).’ Indeed, the authors note that the evidence remains thin, so they were unable to draw any firm conclusions about causal mechanisms.

Moreover, when the word “sanction” was used, it was most commonly referred to as being an *absent* rather than present condition (see Grossman et al., 2017; McNeil and Malena, 2010; Tsai, 2007; for further discussion). According to Waddington et al., (2019: 82), in general, the social sanctions threat of local civic engagement was not deemed to be strong enough to ‘overcome the power difference between providers and interventions often failed to elicit responses in service provider actions, leading to a break in the causal chain.’ This is an issue to which we will return when considering the details of individual studies below.

Many studies cited also did not include an explicit citizen engagement component with service providers or public servants as a form of civil society collective action. A number of the most frequently cited studies in the reviews and wider literature refer to sanctions for corrupt politicians or mayors or exclusively to internal accountability mechanisms related to civil servants or service providers’ performance, but had no clear links to citizen engagement and collective action (e.g. Boas and Hidalgo, 2019; Caldeira, 2012; Chen et al., 2015; Ferraz and Finan, 2008; Raffler, 2019; Wang, 2019; Zheng et al., 2014).<sup>10</sup> Indeed, when examined further, a number of the studies cited as evidence that social or formal sanctions were important did not make such a claim explicit (e.g. Björkman and Svensson, 2009; Björkman et al., 2014; Bradley and Igras 2005; Pandey et al., 2009; Ravallion et al., 2013; Reinikka and Svensson, 2005). While it was possible there was some connection to sanctions this seemed mostly to rely on the interpretation of reviewers rather than the studies themselves. We also found that various other commonly cited reviews did not specify which specific empirical

<sup>10</sup> In Mendiburu’s (2020) review of citizen participation in Supreme Audit Institutions in Argentina, Brazil, Chile, Colombia, Mexico, and Peru, citizen participation was limited to a single phase in the process (simply in sharing audit findings) or where complaints were individual rather than collective (e.g., complaints mechanism).

evidence from which contexts supported claims in relation to sanctions (Agarwal et al., 2009; Joshi, 2010), so it was not always possible to substantiate various claims empirically.

Pressure was a commonly used term, but it was rarely clear *how* such pressure was manifested or contributed to service delivery outcomes. According to McNeil and Malena (2010: 216 – 217), this pressure is argued to manifest itself through ‘the potential for public shaming through media coverage, public forums, and public interface meetings between citizens and state officials,’ including in the Extractive Industries Transparency Initiative (EITI).<sup>11</sup> McNeil and Malena do not present evidence to show how outcomes were achieved through public shaming. The EITI’s most recent theory of change makes the same connection, listing “naming and shaming” as one of their three key models. However, the theory of change notes that the EITI ‘did not equip citizens with tools to demand accountability’ nor did it have the ‘power to prosecute,’ or ‘outline mechanisms that would ensure accountability’ directly (Le Billion et al., 2020: 37). Thus, it is not clear the degree to which such connections are based on evidence.

Finally, as Tsai et al., (2019) note, within the 11 reviews, evidence is highly concentrated in only a handful of countries, most notably, China, Indonesia, and Uganda. So, as Tsai et al., (2019) stress, we should therefore be careful with claims of external validity, despite the generally high quality of the meta-reviews assessed. It is for this reason that we conducted a search for additional studies, and in the following section, we will consider findings across 35 cases with a wider spectrum of contexts. This should allow us to draw potentially more transferable lessons.

## Findings on sanctions from individual studies

### *Use of formal and social sanctions*

Efforts to impose formal or social sanctions would appear to be relatively common practice. Two-thirds of the 35 cases reviewed involved formal sanction threats, just under half included some effort to enforce sanctions, and three-quarters of cases involved some effort to impose social sanctions. Overall, evidence of an attempt to impose either social or formal sanctions was found across 17 countries.<sup>12</sup> Hence, we suggest that efforts to impose different forms of sanctions are likely to be quite widespread, even if they do not feature prominently in past evidence reviews in the transparency and accountability field.

### *Formal sanctions*

Formal sanction threat intensity varied significantly across cases. Some of the formal sanction threats reviewed implied heavy punishments. These reached the height of government lawsuits (Arroyo and Sirker, 2005; Herrera and Mayka, 2019), withdrawal of contracts (Aiyar, 2010; Banerjee et al., 2007; Duflo et al., 2009; Gaduh et al., 2020; Majeed, 2011; Woodhouse, 2005) and pay deductions (Bailey and Mujune, 2021; Banerjee et al. 2007; Duflo and Hanna, 2005; Duflo et al., 2009; Gaduh et al., 2020). Elsewhere, formal sanction threats were implicit, and these featured through efforts such as school inspections (Kremer et al., 2005), or more commonly, sanctions were *assumed* through the threat of unspecified punishments (Beasley and Huillery, 2012; Klugman, 2013; Molina, 2014).

Surveys and administrative reviews suggest that only a small proportion of staff faced formal sanctions in practice (Kremer et al., 2005; MoFALD, 2016; Pande and Dubbudu, 2017). Consistent with findings in research on Grievance Redress Mechanisms (GRMs) in state social protection programs in Argentina, Mexico, and Brazil (Gruenberg and Pereyra Iraola, 2008), the limited data available in India and Nepal, for instance, suggest that formal sanctions tend to be disbursed in less

---

<sup>11</sup> These referred to initiatives in Benin, Malawi, Nigeria, and Uganda.

<sup>12</sup> Bangladesh, Colombia, Dominican Republic, Ethiopia, Ghana, Guatemala, India, Indonesia, Kenya, Macedonia, Mozambique, Nepal, Niger, Peru, Philippines, South Africa, and Uganda



than 5% of cases (Kremer et al., 2005; MoFALD, 2016; Pande and Dubbudu, 2017). For example, Kremer et al., (2005) found that across 20 Indian states, ‘only one head teacher in nearly 3,000 public schools reported ever dismissing a teacher for repeated absence (Kremer et al., 2005: 661 – 662).’<sup>13</sup> The authors thus argued that ‘teachers fe[el] little risk of being fired for absence (Kremer et al., 2005: 661),’ a sentiment which was shared by scholars and evaluators more widely in India and Nepal, as well as in Indonesia (Pande and Dubbudu, 2017; Schjødt, 2017; Woodhouse, 2005). So, while various interventions attempted to trigger formal sanctions, these efforts were rarely successful.

Where formal sanctions were enforced, these were more commonly administrative sanctions imposed at local level from within a particular service (i.e., within managers’ direct control). These included pay deductions (Bailey and Mujune, 2021; Banerjee et al., 2007; Duflo and Hanna, 2005; Duflo et al., 2010; Gaduh et al., 2020), disciplinary measures such as temporary suspension (Duflo et al., 2010; Pande and Dubbudu, 2017; Schjødt, 2017), and relocation of poorly behaved staff (Ayliffe, 2017; Schaaf et al., 2020; Schjødt, 2017). However, we generally found low levels of acceptance by managers to enforce harder punishments. As we will discuss below, various cases demonstrate that intensive monitoring and threats from above were difficult to sustain and their effects were not sustainable (Banerjee et al., 2007; Duflo and Hanna, 2005; Duflo et al., 2009; Majeed, 2011).

Just under a third of cases reviewed were able to trigger administrative or judicial sanctions imposed from higher administrative levels, yet these generally took longer to impose, especially when investigation and prosecution were decoupled. In general, matching recent findings from Mendiburu (2020) and IBP and INTOSAI (2020), we found that connections between civil society efforts and horizontal accountability agencies such as ombudsmen or the courts were generally quite poor, and success in imposing judicial sanctions in practice was highly mixed (Aston, 2015; Feruglio, 2017; Herrera and Mayka, 2019; Pande and Dubbudu, 2017; Schjødt, 2017; Woodhouse, 2005). Imposing such sanctions also appeared to require the legal capacity and autonomy that many low- and middle-income countries typically lack (Herrera and Mayka, 2019).

### *Social sanctions*

For social sanctions, we found it helpful to divide these into “hard” sanctions and “soft” sanctions. For hard sanctions, we found various confrontational efforts which emphasized “naming and shaming” of service providers and civil servants through media exposure, protest, or litigation.

Three-quarters of cases reviewed employed a harder form of sanctions as a deliberate strategy, across a wide variety of contexts. This was most notable in India, Indonesia, Colombia, The Philippines, Mozambique, Macedonia, and Uganda which had more than a single case in the sample. Efforts to name and shame or to litigate were slightly more common at the sub-national and national level than at the local level (Anderson et al., 2020; Aston, 2015; Bailey and Mujune, 2021; Banerjee et al., 2018; Gaduh et al., 2020; Hernández et al., 2017, 2020; Jones et al., 2019; Joshi, 2017; Larsen, 2015; Molina, 2014; Olken, 2007; Poli et al., 2019; Walker, 2018; World Bank, 2020; see also Peiffer et al., 2021 and Tsai, 2021 for examples *without* citizen-driven efforts).

In a quarter of cases, interventions employed a form of soft sanctions as part of their approach. In general, this approach involved leveraging social norms and persuasion (described in figure 3 below) and was not designed to openly confront and damage or destroy the reputations of counterparts. This approach was mostly found in the health sector and was mostly at the local level (Aston, 2015; Walker, 2017; Joshi, 2017; Feruglio, 2017; Gurang et al., 2019; Hernández et al., 2017, 2020; Schaaf et al., 2020). A handful of examples were also found in the social protection sector, education, housing, and contracting (Aceron, 2018; Ayliffe, 2017; Edwards et al., 2020; Feruglio, 2017; Majeed, 2011; Schjødt, 2017). As will be discussed below, there was also sometimes overlap between harder and softer forms.

---

<sup>13</sup> This was 35 out of 600 in private schools, which is suggested as one reason private schools had lower absentee rates.

Drawing on Westhorp et al’s (2014) rendering of community accountability in the education sector, we identified five mechanisms related to the threat or disbursement of formal or social sanctions which are outlined below:

*Figure 3. Sanction-based Mechanisms*

<b>Mechanism</b>	<b>Notable Examples</b>	<b>Cases</b>
<b>Sticks:</b> in which service providers and civil servants respond to the actual application of formal sanctions such as fines, pay deductions, or other disciplinary measures (i.e., response to punishment), who are alerted by citizens	<ul style="list-style-type: none"> <li>• Kecamatan (Subdistrict) Development Project (KDP) (infrastructure, Indonesia)</li> <li>• Improving Teacher Performance and Accountability – KIAT Guru (education, Indonesia)</li> </ul>	17
<b>Big brother is watching:</b> in which service providers and/or civil servants respond in anticipation of the application of formal sanctions from superiors (i.e., fear of being caught and punished), who are alerted by citizens	<ul style="list-style-type: none"> <li>• Philippines Textbook Program (education, The Philippines)</li> <li>• Seva Mandir (education, India)</li> </ul>	15
<b>Diagonal accountability:</b> in which service providers and civil servants respond to the threat or imposition of formal sanctions by horizontal accountability agencies such as courts or Supreme Audit Institutions (SAIs), which responded to CSOs’ complaints (i.e., response to investigation findings)	<ul style="list-style-type: none"> <li>• Social audits in Mahatma Gandhi National Rural Employment Guarantee Scheme –MGNREGS (social protection, India)</li> <li>• Local Governance and Community Development Program – LGCDP (social protection, Nepal)</li> </ul>	10
<b>Litigation:</b> in which service providers or civil servants respond to legal investigations or lawsuits supported by community paralegals and/or legal aid organizations (i.e., response to investigation findings), prompted or led by civil society efforts	<ul style="list-style-type: none"> <li>• Human Resources for Health (HRH) campaign, (health, Uganda)</li> <li>• Community Defenders of the Right to Health – REDC-Salud (health, Guatemala)</li> </ul>	11
<b>Naming and shaming:</b> in which service providers and civil servants respond to citizen efforts to name and shame them and change their behavior as a result (i.e., response to a sense of shame)	<ul style="list-style-type: none"> <li>• Human Resources for Health (HRH) campaign (health, Uganda)</li> <li>• I am Aware Project (multi-sector, Ghana)</li> </ul>	15

**Source:** own construction

These mechanisms were often found in combination with one another or with complementary mechanisms that did not explicitly emphasize sanctions. We found that sanction-based mechanisms were most often supported by civil society data collection and monitoring of service delivery. Civil society actors would then pass this information on to state actors which had the authority to follow up (what Westhorp et al., 2014 refer to as an “eyes and ears” mechanism).<sup>14</sup> We also found a mechanism which relied on a much softer form of pressure and which may or may not entail sanctions, whereby civil society actors collaborated with service providers or civil servants and persuaded them through an appeal to shared norms and values (i.e., social norms). We labeled this “power of persuasion.” As this softer form of approach does not fit with how sanctions are typically characterized in the literature, it will be discussed in the section below on the potential for synergies between confrontational and collaborative approaches.

<sup>14</sup> This was found in 8/10 cases.

### ***How effective were initiatives in contributing to impact?***

In half of the cases reviewed, we were able to establish a likely link between social or formal sanctions and intermediate effects. It should be noted that not all mentions that sanctions were significant offered evidence to demonstrate the link to outcomes. However, half of the cases could be said to demonstrate such a link to some moderate degree of confidence. Effects that were likely linked to sanctions included increased service provider awareness and motivation (Gurang et al., 2019), increased user satisfaction (Molina, 2014), reduced teacher and nurse absenteeism (Bailey and Mujune, 2021; Banerjee et al., 2007; Duflo and Hanna, 2005), increased instruction time (Duflo et al., 2010; Gaduh et al., 2020), increased availability of quality textbooks (Arroyo and Sirker, 2005; Majeed, 2011), receipt of medicines (Herrera and Mayka, 2019), lower levels of missing expenditures (Olken, 2007), improved infrastructure quality (Beasley and Huillery, 2012), increased funding and staffing made available (Buntaine and Daniels, 2020; Larsen, 2015; Poli et al., 2019), and parents increasing investments in their children's education (Gaduh et al., 2020). These examples demonstrate that both social and formal sanctions have the potential to contribute to outcomes in various aspects of service delivery, at least over the short term. At the impact level, we found documented evidence in only four of the cases reviewed. These included improvements in students' test scores in language and mathematics and the increased provision of a food subsidy (Banerjee et al., 2018; Duflo and Hanna, 2005; Duflo et al., 2009; Gaduh et al., 2020).

When comparing financial, legal, and employment-related punishments, it may be argued that the threat or imposition of financial punishments (e.g. pay deductions or loss of contracts) may have been more effective disincentives than the less credible and more contingent threat of being fired or the likelihood of criminal conviction (Bailey and Mujune, 2021; Banerjee et al., 2007; Duflo and Hanna, 2005; Duflo et al., 2009; Duflo et al., 2010; Gaduh et al., 2020; Kremer et al., 2005; Majeed, 2011; Pande and Dubbudu, 2017; Schjødt, 2017). However, it should be noted that financial punishments met with substantial resistance and in some cases had serious ethical drawbacks, as will be discussed below under negative effects.

Actual punishment is not always necessary to change behavior. In several cases, the credible *threat* of punishment appears to have sufficed. Yet, as the case of The Philippines Textbook Count in box 1 below shows, success may rely on propitious circumstances, as well as intensive civil society monitoring.

### **Box 1. Watching carefully may suffice, in good times**

The Philippines Textbook Program demonstrates how financial (dis)incentives and intensive monitoring may provide an effective deterrence for potential corruption in contracting in the education sector, with a propitious enabling environment. Public attention to corruption was heightened by scandals which led to the ousting of President Joseph Estrada in 2001 (Aceron, 2018). The World Bank had outlined an anti-corruption strategy in 2000 and approved a loan under the Second Social Expenditure Management Project which included funds for the distribution of 37 million textbooks to public schools and mandated international competitive bidding as a loan condition. The program was also supported by the passing of the Government Procurement Reform act in 2003. Government Watch (G-Watch) and Procurement Watch were engaged to monitor the Education Department's bidding process, including deliberations of the department's Bids and Awards Committee (and videotaping proceedings). The National Citizens' Movement for Free Elections (Namfrel), and later the Scouts, then inspected the quality of textbooks and tracked deliveries. The Education Department would not pay for deliveries until the monitoring data were received and any rejected shipments had to be rectified at the publisher's expense (Majeed, 2011: 10).

Procurement consultant and former public-school teacher Aida Carpennero noted that: "we would use a carrot-and-stick approach depending on the employee group and how intransigent or cooperative they were (Majeed, 2011: 6)." However, with no evidence of sanctions presented in the case, as Carpennero points out, "sending out signals" and "watching carefully" was deemed "sufficient (*ibid*)." As a result, according to the Instructional Materials Committee Secretariat, every student had a book in the four main subjects of English, mathematics, science, and social studies. The initiative contributed to reducing the price of textbooks, improved their physical quality and shortened the procurement cycle significantly (Aceron, 2018). However, in 2005 a highly supportive official, Juan Miguel Luz, was a casualty of government reshuffles, two years later G-Watch handed over to the Department of Education and donor funding was pulled. It was thus broadly argued that things went "back to the previous practice that suppliers deliver on their own time and schedules (Majeed, 2011: 14, see also Aceron, 2018)." With President Gloria Macapagal-Arroyo facing accusations of electoral fraud and large-scale corruption, Joy Aceron (2018: 91) suggests that a 'large anti-corruption drive like [the] Textbook Count may not have been credible or even feasible.' While this perspective may perhaps be overly pessimistic, the case does raise questions regarding how well a "big brother" is watching mechanism may work in more troubled times.

Confrontation tends to happen during perceived windows of opportunity. While success was achieved, for example, in The Philippines Textbook Count with a carefully handled media engagement strategy and with broadly "constructive engagement" with government (Aceron, 2018: 76, 85), we found many cases which promoted more confrontational approaches to advocacy, particularly at sub-national and national levels. This approach was pursued in contexts as diverse as the Dominican Republic, Ghana, Guatemala, Pakistan, Peru, The Philippines, Indonesia, Macedonia, Mozambique, Nepal, South Africa, and Uganda (Anderson et al., 2020; Aston, 2015; Bailey and Mujune, 2021; Banerjee et al., 2018; Edwards et al., 2020; Feruglio, 2017; Jones et al., 2019; Joshi, 2017; Olken, 2007; Poli et al., 2019; Schjødt, 2017).

Rather than a *modus operandi*, in many cases, confrontational strategies were employed in short bursts, particularly during election or budgeting cycles (i.e., windows of opportunity). Further discussion will be found in the section on confrontation and collaboration.

Social sanctions through "naming and shaming" can be an important strategy to influence policies or budgets, even if these are not always implemented. Just over a tenth of cases reviewed were able to achieve policy or budget outcomes. In large part, this was because a relatively smaller proportion of the interventions reviewed set such changes as explicit goals. Where national-level policy or budget changes took place (Uganda, Dominican Republic, and Kenya), this was achieved through multi-stakeholder campaigns of substantial (and often unprecedented) scale with dozens or even hundreds of organizations forming part of coalitions and where the advocacy issue (e.g., investment in

education, human resources in health) was exceptionally politically salient (Feruglio, 2017; Joshi, 2017; Larsen, 2015; Poli et al., 2019; Walker, 2018). For sub-national policy or budget change (Ghana and Mozambique), “naming and shaming” through media exposure was argued to play an important role (Anderson et al., 2020; Jones et al., 2019).

“Naming and shaming” was, however, a hit-and-miss strategy. We found evidence of success in only a quarter of cases in which this was a potential mechanism. Media exposure was widely argued to be important in cases reviewed (Anderson et al., 2020; Feruglio, 2017; Jones et al., 2019; Joshi, 2017; Schjødt, 2017; Woodhouse, 2005). However, very rarely was the evidence for *how* media made a difference compelling. For instance, despite Anderson et al’s (2020) argument regarding the importance of “naming and shaming” through the media in Pakistan, the evaluation of the program – Empowerment, Voice and Accountability for Better Health and Nutrition – suggests that the presence of media in Provincial Advocacy Forums ‘may be the reason why government [was] *less* interested in attending (Taylor and Khan, 2017: 14, our emphasis).’ With reporters or cameramen constituting a quarter of participants in these forums, it is perhaps unsurprising that government attendance declined over time.

However, testimonial evidence from public servants interviewed in Ghana demonstrates that media exposure *can* put low-performing public servants under pressure to ‘rise from their slumber’ and ‘puts [them] on [their] toes to work harder to improve services (Jones et al., 2019: 26).’ In this case, the I am Aware project benchmarked districts against one another through a league table and media used the league table to name and shame low performing districts through sensationalized reporting. The evaluation suggests that newer and lower-capacity districts may have been particularly susceptible to the pressure of media exposure (Jones et al., 2019). One more dramatic example of the benefits and costs of naming and shaming at scale can be found in the case of the Human Rights for Health (HRH) Campaign which was coordinated by the White Ribbon Alliance and World Vision in Uganda which is detailed in box 2 below.

## Box 2. Breakthroughs and blowback of campaigns at scale

The Human Resources for Health (HRH) campaign brought together the “Petition 16” campaign, the Coalition to End Maternal Mortality and the Coalition for Maternal, Child and New-born Health into a mega-coalition. Since early 2011, the Ministry of Health and parliamentary Social Services Committee (SSC) had strengthened links with CSOs in USAID-funded “Uganda Capacity Plus” programme. The SSC asked CSOs for support in scrutinising the Ministry of Health’s budget. Yet, by late 2011, CSOs pivoted from what Jillian Larson (2015: 16) terms a “gentleman’s approach” to one which the Ministry of Health considered an attempt at ‘sabotage.’ While MPs on the Social Services Committee appreciated CSOs’ analysis, Larson (2015: 16) notes that the Ministry ‘felt that it was under attack by CSOs [and that they had] used their position in working groups and committees to steal insider information which was then brought before parliament.’

Following government commitments made in the Inter-Parliamentary Union (IPU) meeting in March 2012, World Vision convened a national dialogue with key decision-makers in June, where policymakers agreed to increase health sector funding. The “Petition 16” to Stop Maternal Mortality conducted peaceful demonstrations outside the Supreme Court for their appeal on the 22<sup>nd</sup> May 2012 which included shaming of government officials for supposedly having received UGX 377 billion of healthcare abroad, which exceeded requested increases in health spending (Larson, 2015: 24; see also Templeton Dunn *et al.* 2017 on how the case was reopened by the Constitutional Court in 2016). Citizens sent Short Message Service (SMS) to parliamentarians saying: “We are watching you: Refuse to pass the budget unless it includes the increase you promised (Walker, 2017: 10).” National television stations and newspapers then went to visit rural clinics. David Walker (2017: 10) argues that ‘senior government officials shamed by revelations that public funding they received to obtain healthcare overseas was on par with the increased health budget requested by campaigners.’ An increased budget was passed which allowed the government to retain and recruit a thousand additional health workers (Republic of Uganda, 2013). However, some CSOs saw this success as a ‘pyrrhic victory’ because their role ‘declined as a result of changes carried out by the executive immediately following the campaign’s budget victory,’ with reduced transparency from cabinet as a result (Larson, 2015: 37).

## *Sustainability and systemic effects*

Sustaining sanctions-based approaches is challenging, and effects appear to be rarely sustainable. For many of the cases reviewed, specific data on sustainability were unavailable. However, where the question was asked, sustainability was a perennial issue. A quarter of these cases showed that interventions which focused on applying social or formal sanctions were not sustainable, and in some cases, effects dissipated within only a few months (Banerjee *et al.*, 2007; Beasley and Huillery, 2012; Duflo and Hanna, 2005; Duflo *et al.*, 2009; Duflo *et al.*, 2010; Fox, 2020). More severe punishments were also more difficult to disburse and enforce. We found this within administrative units which were asked to enforce administrative punishments such as in India’s education and health sectors. Staff blocked and sabotaged efforts (Banerjee *et al.*, 2007; Duflo and Hanna, 2005).

In India’s health sector for instance, Seva Mandir found that while the intervention significantly reduced nurses’ absenteeism for a few months, by the end of the intervention, the attendance of assistant midwives was actually *lower* than the year before the intervention and treatment assistant nurse midwives were actually *less* likely to be present than control assistant nurse midwives (Banerjee *et al.*, 2007). This lack of sustainability thus suggests that positive effects were unlikely to be systemic. As the cases of the Poverty Safety Net Program (PSNP) in Ethiopia and the Child-Sensitive Social Protection Program (CSSP) in Nepal suggest, when sanctions had been disbursed, it is also generally unclear what wider deterrence effect they may have had on the misbehavior of other actors, or even for some of the same actors that were punished. In both cases, punishment appeared merely to relocate problems from one area to another as nurses were simply moved from one location to another (Aslam, 2017; Schøjdt, 2017). In Indonesia’s KDP it was initially suggested that there may be such a deterrence effect to other locations through the threat of withdrawing donor funds (Woodhouse, 2005). While it is unclear that these deterrence effects were sustained over time after the program

ended, certain aspects of the KDP later found expression through the Village Law of 2014 (Fox, 2020).

Overall, it seems likely that, had more research or evaluations assessed sustainability, an even higher proportion would likely be deemed unsustainable or may yet be unsustainable, given the level of service provider and manager resistance on the grounds of perceived unfairness (Banerjee et al., 2007; Gaduh et al., 2020). This lack of sustainability was by no means confined to “principal-agent” inspired experiments or state administrative systems with links to citizen engagement efforts.<sup>15</sup> The cases of the HRH campaign in Uganda and the Coalition for Dignified Education in the Dominican Republic both suggest that coalitional advocacy efforts which initially contributed to budget changes seriously struggled to make government commitments stick even just months after initial breakthroughs (Bailey and Mujune, 2021; Larson, 2015; Poli et al., 2019).

### *Negative effects*

Efforts to impose sanctions may have significant unintended costs. In just under half of cases reviewed, we also found negative effects, and in the vast majority of these cases, negative effects were directly linked to sanctioning efforts.<sup>16</sup> These effects included discrediting, threatened loss of funding (Feruglio, 2017), reprisals for advocates and whistle-blowers, including court cases, and staff being relocated for speaking up (Aslam, 2017; Hernández et al., 2020; Klugman, 2013), the denial of access to services (Bailey and Mujune, 2021), an increased concentration of poorly performing staff in some areas (Ayliffe, 2017; Schøjt, 2017), and even threats of violence to MPs and cabinet reshuffles for those who collaborated with campaigners (Larson, 2015).

Imposing both social and formal sanctions can be counterproductive. Some interventions were said to have damaged staff morale, generating conflict among staff and between staff and community members and between staff and patients, damaging trust (Gaduh et al., 2020; World Bank, 2020, see also Lodenstein et al., 2018; Peiffer et al., 2021). Some sanctioning efforts decreased organizational leverage with government counterparts and caused accountability operations to temporarily shut down (Aston, 2018; Edwards et al., 2020) and decreased budget or contracting transparency (Bailey and Mujune, 2021; Edwards et al., 2020; Larson, 2015).

Other interventions which appear to be more inspired by a “principal-agent” approach to accountability and typically disbursed administrative sanctions also had significant negative effects such as a decrease in the attendance of some staff (Banerjee et al., 2007; Gaduh et al., 2020).<sup>17</sup> Given the high likelihood of under-reporting negative outcomes, it seems that the proportion of negative effects is almost certainly higher than we found in this study. While there is little evidence of citizen-driven efforts, the establishment of the Health Monitoring Unit (HMU) in Uganda offers a similar view of undermining staff morale and citizens’ trust in efforts to tackle bribery (Peiffer et al., 2021). Reflecting Mansbridge’s (2014) caution regarding principal-agent approaches, the case of the Improving Teacher Performance and Accountability (KIAT Guru) program in box 3 below shows, when either formal or social sanctions are employed as a key part of an intervention’s strategy, the costs of achieving service delivery outcomes might even be outweighed by the social cost of repercussions within communities.

---

<sup>15</sup> “Principals” are citizens and “agents” are service providers or bureaucrats in this understanding (see Gaduh et al., 2019: 1 and Raffler et al., 2019).

<sup>16</sup> This is slightly higher than Gaventa and Barrett’s (2010: 51) finding that 25% of outcomes from citizen engagement interventions found negative outcomes. Yet, unlike Gaventa and Barrett, we found that backlash and reprisals were more likely from what they term contentious (i.e., confrontational) approaches than from formal participatory governance spaces. Indeed, our findings are consistent with recent findings that 62% of protests globally between 2006-2011 resulted in arrests, surveillance, injuries, and deaths (Ortiz et al., 2022). Overall, it seems likely that more contentious efforts have more negative effects in response.

<sup>17</sup> As Peiffer et al., note (2021), meta-analyses of corruption studies provide limited evidence of the effectiveness of most principal-agent theory-inspired interventions.

### **Box 3. Surface benefits and hidden costs of punishment**

The Improving Teacher Performance and Accountability (KIAT Guru) program in rural Indonesia is one recent example which shows a sticks mechanism in operation, and that making teachers' less comfortable can contribute to some improvements in learning outcomes. Building on the experience of Seva Mandir in rural India, we find that using cameras to monitor teacher attendance and enabling user committees to determine the level of the allowance for relocated teachers prompted teachers to be more diligent and disciplined, as they spent more time in the classroom teaching, and this, in turn, led to improved test scores in language and mathematics (Gaduh et al., 2020).

While over ¾ of teachers had a written justification for their absence, the World Bank's (2020) qualitative assessment suggests that most teachers were put under pressure. Teachers' fear of losing their allowance was a clear motivation. Some teachers came to school even when they were sick for fear their pay would be reduced. Yet, the initiative damaged staff morale and created conflict among teachers and between teachers and community members. As only 35% of teachers qualified for the teachers' allowance in the first place, the intervention generated backlash from teachers who were ineligible. Some teachers went on strike because of this, others decreased their attendance in the classroom. As user committees did not know what teachers should be scored on, but set teachers allowance levels anyway, this created further conflict. Teachers also dropped creative components of lessons because these were not in the service agreement (World Bank, 2020). As in UNESCO's (2017) global assessment, it might be argued that such efforts to rapidly maximize test score improvement via sanctions may have undermined overall education quality and more important forms of learning.

### ***Different sanctions for different stakeholders***

Some actors appear to be the preferred targets of sanctions. We found that three broad types of stakeholders were the most common targets of punishment: absentee nurses or teachers who had their pay or allowances reduced (Bailey and Mujune, 2021; Banerjee et al., 2007; Duflo and Hanna, 2005; Duflo et al., 2009; Gaduh et al., 2020), offending officials who were either suspended, relocated, or fired (Ayliffe, 2017; Anderson et al., 2020; Pande and Dubbudu, 2017; Schøjdt, 2017; Woodhouse, 2005), and contractors which had to cover the cost of rejected materials (Majeed, 2011) or faced lawsuits, alongside civil servants who were involved in contracting (Arroyo and Sirker, 2005; Woodhouse, 2005; see also Coelho, 2018). It does not appear that this was because the transgressions of these actors were easier to observe or that they behaved significantly worse than other members of staff.

Closer relationships may perhaps deter stakeholders' appetite to impose sanctions. In the health sector, we found a notable preference for soft sanctions over hard sanctions. In part, this has to do with the nature of relationships between patients and clinical staff. Where there are significant asymmetries in knowledge, and where relationships are intimate (e.g., doctor-patient, familial or kinship relations), service users seem more comfortable applying soft sanctions than hard sanctions. As we saw in Peru, Nepal, Niger, and Mozambique, closer relationships created disincentives for confrontation, and a "policing" approach to monitoring was also deemed inappropriate and counterproductive. Instead, we found an endorsement of 'constructive' and 'partnership' approaches which 'leveraged shared values and norms (Aston, 2015; Beasley and Huillery, 2012; Feinglass, et al., 2016; Gurang et al., 2019: 1; Schaaf et al., 2020: 9, see also Arkedis et al., 2019 and World Bank, 2020).' Fear of reprisals was also an important factor dissuading actors from pursuing a confrontational approach to social sanctions in more dominant and hierarchical political systems such as in Ethiopia (Aslam, 2017).

Actors may, instead, prefer to sanction "others." We found a different approach when the actor/organization targeted was perceived to be outside the group. This is most notable in cases where



the community and state apparatus both appear to have incentives to seek harder sanctions on short term consultants, contractors, and suppliers. We found this across sectors and geographies. In India's health sector, we saw this with contract nurses (Banerjee et al., 2007), in the education sector we saw this with contract teachers in Kenya (Duflo et al., 2009) and teachers with a relocation allowance in Indonesia (Gaduh et al., 2020),<sup>18</sup> as well as with firms contracted to deliver textbooks in Bangladesh and The Philippines (Arroyo and Sirker, 2005; Majeed, 2011). We found a similar example for government contracted staff in the social protection sector in India (Aiyar, 2010), as well as in building infrastructure in Uganda and Colombia (Buntaine and Daniels, 2020; Molina, 2014; see also Coelho, 2018 in Brazil). Bypassing contractors was even an explicit strategy employed to avoid corruption in Indonesia's KDP (Woodhouse, 2005). We thus found that, both service users and providers are perhaps more likely to pursue harder sanctions with stakeholders with whom they have a more ephemeral relationship.

### ***Collaboration and/or confrontation?***

The previous section offered several case studies which illustrate that different sanction-based mechanisms can be combined. However, it has also been argued in the scholarly literature that there may be productive combinations of collaborative and confrontational tactics – i.e., hybrids (see Anderson et al., 2020; Bukenya et al., 2012; Fox, 2014, 2016; Joshi, 2017). In our cases, this would entail a productive combination of sanctions-based mechanisms and what we labeled the “power of persuasion.”

Various cases reviewed illustrated a productive relationship between civil society and state actors. Most notable among these were Civil Society Forum in Health's (*Foro Salud*) long-standing partnership with the Ombudsman in Peru's health sector, Government Watch's (G-Watch) strong relationship with senior figures within the Department of Education in The Philippines, parliamentary actors linked to the Human Rights for Health campaign in Uganda, and with the judiciary in Colombia (Aston, 2015, 2016; Herrera and Mayka, 2019; Larsen, 2015; Majeed, 2011). In Peru, The Philippines, Uganda, and Colombia, state actors were able to provide tangible support or opened spaces which clearly contributed to outcomes. However, these cases generally did not combine collaborative and confrontational tactics in time and space (as per Fox, 2016). Where they did, for example in Uganda's HRH campaign, we found that some positive state relationships were sacrificed for others (Ministry of Health vs. Members of Parliament – MPs) (see box 2). Indeed, we found cases such as that of Open Up Contracting in The Philippines in which the “naming and shaming” of senior politicians by The Philippines Centre for Investigative Journalism nearly derailed the collaborative efforts of other actors to persuade The Philippines Government Electronic Procurement System and the Commission of Audit to accept recommendations to amend an Open Government Partnership action plan (Edwards et al., 2020). This, therefore, suggests that targeted conflict may have a substantial risk of backfiring.

Collaborative and confrontational tactics often tend to happen in different locations, even if these are within the same broad strategy. Perhaps, in part, due to the aforementioned risks, a number of cases reviewed employed different tactics in different locations but at the same level of government and/or different tactics at different levels of government (Anderson et al., 2020; Flores and Hernández, 2018; Joshi, 2017; Larson, 2015; Walker, 2017). In Guatemala's health sector, for example, Joshi (2017) notes that civil society actors employed collaborative tactics locally and more confrontational tactics nationally. In fact, the Centre for the Study of Equity and Governance in Health Systems (CEGSS) even developed a decision tree to support Community Defenders of the Right to Health – *REDC-Salud* to determine when to collaborate and when to confront authorities. For instance, we found a collaborative strategy in San Marcos which had supportive authorities and the use of petitions and litigation in Totonicapán and Santa Maria with less supportive authorities (Joshi 2017; Hernández et

---

<sup>18</sup> It is worth noting that this short-termism with precarious labor contracts appears to, in part, be a feature of experimental designs, which appear to skew initiatives towards harder sanctions, which absent these evaluation designs might not be chosen with the same (perhaps misleading) degree of frequency.

al., 2019). In Uganda, on the other hand, we found local and national level collaborative efforts, and then efforts at national level became more confrontational during the HRH campaign (see box 2). Like Anderson et al's (2020: 32) findings from DFID's accountability programming, we were unable to find any compelling evidence that confrontational tactics enabled broader *collaboration* within government and, in turn, lead to more effective policy implementation.

Moreover, we found that horizontal accountability agencies generally offer weak protection, and typically have limited leverage to pressurize other government actors. In Guatemala, when authorities refused to commit to a resolution, the aim was to seek mediation from the Ombudsman's office, and potentially naming and shaming through press conferences (Flores and Hernández, 2018). Disappointing returns from efforts to trigger legal redress<sup>19</sup> prompted members of *REDC-Salud* to present petitions and collective grievances to various governmental oversight agencies or even parliamentarians.<sup>20</sup> These too were ineffective in triggering sanctions to remove health authorities, and filing collective grievances generally failed to stem mistreatment (Hernández et al., 2020). Worse, the more confrontational tactics employed in some areas locally even resulted in reprisals through a lawsuit against one community healthcare defender for her alleged threats and intimidation (Batzin et al., 2020), and we find more violent reprisals in social auditing efforts in India (see Goetz and Jenkins, 2005).

The repercussions in the Uganda case further show that even with the support of senior politicians and a parliamentary committee, efforts were still *perceived* as being an attempt at "sabotage" by other state actors, and both supportive state actors and civil society actors were punished as a result (Larsen, 2015).<sup>21</sup> Rarely was support from higher-level staff, horizontal accountability agencies, or parliamentarians an effective form of protection or mitigation, nor was it clear they offered any substantial leverage (see Aston, 2015; Bailey and Mujune, 2021; Batzin et al., 2020; Fox, 2014: 32, 2016). Only in what are considered extreme cases in Colombia do we find any real evidence of leverage from oversight agencies or the judiciary. It was, in fact, state entities which decided to pool legal petitions (*acciones populares*) and writs of protection (*tutelas*) and mandated various oversight mechanisms such as public hearings (*audiencias públicas*) and citizen oversight committees (*comité de verificación*) to enforce environmental protections and constitutionally guaranteed provisions in the health sector (Herrera and Mayka, 2019). Yet, as this is considered by the authors as an extreme case, we should be careful in assessing the potential transferability of findings.

We found that confrontational and collaborative approaches were often not employed concurrently. We agree with Feruglio (2017: 25) that there are important 'chronological considerations when choosing one approach or another.' From our review of the evidence, this temporal dimension matters greatly, yet not merely in terms of employing confrontation as a "last resort." In several cases reviewed, we found *periods* of collaboration and periods of confrontation were employed, with varying degrees of effectiveness. In Peru, for example, one ombudsman's representative in Puno noted that "at first one could feel the rejection of [staff]... they are not accustomed to external oversight from civil society. [Now] the doctors have said that they can see it from another perspective, because before they only presented the negative cases and now, they feel recognized, and this motivates them (Aston, 2015: 5)." Service providers saw the dialogues *Foro Salud*, the ombudsmen and CARE facilitated as an opportunity to explain to patients why, from their perspective, problems arose in the health facilities, and this generated incentive for them to collaborate (Vargas, 2013). Indeed, at national level, as new leadership in *Foro Salud* pursued a more confrontational approach in the healthcare reform of 2013 this prompted ministerial aides to question the very representativeness of the forum and consulted other civil society groups instead (Aston, 2018: 4 – 5).

---

<sup>19</sup> Of the 818 complaints registered between February and September 2015, none of these were accelerated for legal redress (CEGSS, 2015). No data was publicly available beyond this date.

<sup>20</sup> Commission against Discrimination and Racism (CODISRA), Defender of Indigenous Women (DEMI), Presidential Commission on Human Rights (COPREDESH) and the Ombudsman.

<sup>21</sup> See also Edwards et al., (2020) for an example of negative spill-over effects from the confrontational tactics of one organization to another which pursued a more collaborative approach.

Redeploying confrontational and collaborative approaches is difficult. *Foro Salud*'s challenges resemble Feruglio's (2017) explanation of Ndifuna Ukwazi and the Social Justice Coalition's frustrated efforts to redeploy a generally confrontational approach to a collaborative one in participatory budgeting in Cape Town, South Africa. Feruglio notes how employing a confrontational approach, at least temporarily, 'closed the door to the possibility of non-confrontational engagement (Feruglio, 2017: 30, see also Overy, 2013).' Each of these cases suggests that pivots from broadly confrontational to collaborative approaches (or vice versa) may be more difficult than is often anticipated. Figure 4 below illustrates several cases where we found strategic pivots at different moments in time (T1, T2, T3), with varying degrees of success:

Figure 4. Redeploying Confrontation and Collaboration

Intervention(s)	T1	T2	T3
<b>Human Resources for Health Campaign (Uganda)</b>	Persuasion + eyes & ears (2009 – 2011) <sup>22</sup>	Litigation + naming & shaming (2012)	Eyes and ears, sticks, + (attempted) naming & shaming (2014 – 2018) <sup>23</sup>
<b>Ndifuna Ukwazi &amp; Social Justice Coalition (South Africa)</b>	Litigation, eyes & ears, + naming & shaming (2012 – 2013)	Naming & shaming & (attempted) persuasion (2013 –)	
<b>Foro Salud citizen monitoring (Peru)</b>	Eyes & ears, (attempted) naming & shaming +(attempted) diagonal accountability (~ 2007)	Persuasion, eyes & ears, + (attempted) diagonal accountability (2008 – 2013) <sup>24</sup>	Eyes & ears, (attempted) naming & shaming, + (attempted) diagonal accountability (2013 – 2014)
<b>Coalition for Education with Dignity (Dominican Republic)</b>	Litigation + naming & shaming (2010 – 2013)	Eyes & ears + (attempted) naming & shaming (2014 – 2016) <sup>25</sup>	Eyes & ears + persuasion (2016 –) <sup>26</sup>

Source: own construction

Ndifuna Ukwazi and *Foro Salud*'s experience suggest that both the intention to confront (i.e., to punish or damage reputations) and the perception of being confrontational by others may lead to path dependent effects. While it may be true that collaborative and confrontational strategies are rarely pure archetypes (see Fox, 2016), overall, the cases in this study raise questions regarding how productive (or not) synergies between confrontational and collaborative approaches may be in various

<sup>22</sup> World Vision's Citizen Voice and Action work prior to the HRH campaign.

<sup>23</sup> This is formally a separate intervention but focused on the same issue of human resourcing. GOAL's ACT Health program and two phases 2014 – 2016 (phase 1 which included an eyes & ears mechanism) and 2016 – 2018 (phase 2). Phase 2 incorporated advocacy campaigns focused on human resourcing which placed greater emphasis on applying pressure and prosecuting managers with threats to reach 'national offices to request investigation and oversight of recalcitrant district-level official (Bailey and Mujune, 2021: 75).'

<sup>24</sup> *Foro Salud* was supported by CARE's Participatory Voices project (2014). There was also a brief period of naming and shaming in Puno around the 2011 elections, which temporarily forced project activities to be shut down.

<sup>25</sup> Oxfam's *Vigilantes* program, which changed its approach after mid-term review in 2015.

<sup>26</sup> Government's How is My School Doing initiative and World Vision's Community Participation in How is My School Doing project.

contexts, and indeed recall questions regarding how easy or difficult it may be to redeploy each approach (see Fung and Wright, 2003: 267, 282).

We found that soft power can, in fact, be powerful. Waddington et al.'s (2019) systematic review argued that 'monitoring interventions that rely on the "soft" power of social sanctions may be more effective when they focus on identifying mutually empowering "win-win" opportunities and ways for citizens and service providers to work together (Waddington et al., 2019: 67).' Yet, the foregoing discussion raises questions regarding differences between the form and severity of sanctions confrontational and collaborative tactics promote. A quarter of cases reviewed fit the more collaborative mold at least for a period or in particular locations (Aceron, 2018; Aston, 2015; Edwards et al., 2020; Gurung et al., 2019; Joshi, 2017; Schaaf et al., 2020; Schøjdt, 2017; Vargas, 2013; Walker, 2017). While civil society actors may perhaps have been able to trigger formal sanctions in some cases, their approach was viewed by service providers and civil servants as being constructive and helped them to better 'understand whether [they were] really providing [adequate] services or not (Gurang et al., 2019: 5)' and as the 'best way to deal with...blame (Gurang et al., 2019: 7),' as one case of social audits in Nepal illustrates. Similarly, Namati's efforts in Mozambique to train community paralegals in the health sector implored that paralegals should "not to go in there and act as if you were police, or go in to judge, no, you go in as if you were working with them, seek to know what it is like there (Schaaf et al., 2020: 9)" and "act [...] in good faith" as part of a "constructive working relationship (*ibid*; see also Feinglass, 2016: 237 on 'amicable resolution')." As a result, health care providers saw paralegals as complementary support, because they were able to raise issues and solve problems that providers were unable to solve on their own.

Several cases revealed forms of legal empowerment and diagonal accountability which were discernibly constructive. The efforts of Hakijamii in Kenya, *Foro Salud* and the Ombudsman in Peru, Namati in Mozambique, and the judiciary in Colombia focused more on legal enforcement and remedial actions than they did on sanctions (Aston, 2015; Feinglass, et al., 2016; Feruglio, 2017; Herrera and Mayka, 2019; Schaaf et al., 2020; Vargas, 2013). Each of these efforts focused on the rule of law and enforcement of legal entitlements, with state agencies often supporting civil society actors, but did not tend to focus on sanctions as a priority. Instead, it could be said that sanctions were merely a consequence of last resort that rarely materialized in practice. As these more constructive approaches had a higher success rate than more adversarial approaches in our sample and had fewer negative effects, it seems important to disentangle legal enforcement and sanctions, or to unbundle the influential concept of "teeth (see also Feruglio, 2017; Fox, 2014)."

In his expansive definition of teeth, Fox (2014: 8, 23, 28) refers to state and public sector responsiveness, state capacity to respond to voice, institutional capacity for accountability (i.e., to apply *negative* legal or administrative sanctions), investigating and verifying complaints and grievances, changing public sector incentive structures to discourage abusive or wasteful behavior, deploying preventative measures to reduce opportunities for corruption or abuse, and even *positive* incentives for following citizen recommendations, as well as pressure from below generated by protest (i.e., voice). It is therefore unsurprising that scholars and practitioners have interpreted the concept in substantially different ways.

Soft peer and reputational pressure can also be effective. These more constructive and collaborative tactics thus better reflect Lodenstein et al.'s (2018) account of informal strategies in Malawi's health sector, whereby community members employed what they term "soft power" to persuade health staff through an appeal to collective values rather than to coerce through "hard" sanctions. The aforementioned cases also partly reflect the language from Lily Tsai's (2007: 13–15, 112, 141) study of accountability in rural China; actors are not 'watchdogs' prompting an 'adversarial relationship,' but instead, 'moral authority' was seen as a source of 'soft power,' leveraging a 'habit of cooperation' and tapping into authorities' and officials' pursuit of 'higher moral standing, sense of 'self-respect,' and opportunities for potential 'personal gain.' Tsai still refers to 'group sanctions,' and that

community members may ‘*implicitly* threaten village officials (2007: 112, 141, our emphasis).’<sup>27</sup> However, rather than being motivated by a fear of being caught and punished or a sense of shame from the public exposure of transgression, Tsai argues that the motivation was commonly to ‘obtain,’ ‘secure’ or ‘maintain face.’<sup>28</sup>

While there is potential in these cases that service providers or bureaucrats may diminish their prestige if they do not live up to expectations, this is not a zero-sum and adversarial contest as we commonly find in the other mechanisms discussed. Instead, in such a mechanism, responsiveness (and service delivery) is primarily achieved by the “soft power” of deliberation, persuasion, and an appeal to mutual interests. While there is reference to the possibility of sanctions, these cases jar with Schedler’s (1999) conceptualization. And thus, what we have termed as “soft sanctions” seems to better fit with deliberative approaches such as collaborative social accountability (Guerzovich and Poli, 2020; Mansbridge, 2014; see also Fung, 2020).

We found that harder forms of accountability are not necessarily more effective. Even though case sampling for this study was primarily designed to find successful cases of harder sanctions, the success rate for initiatives in the sample which had periods which pursued a more collaborative approach through persuasion was slightly higher than those which pursued more confrontational strategies and harder sanctions, and these initiatives tended also to have fewer negative effects. Indeed, more constructive approaches to legal empowerment and diagonal accountability reviewed were actually more successful. Furthermore, a brief comparison with interventions in broadly similar contexts to those within the sample – Uganda, Ghana, Malawi, Peru, and India – suggest that comparable positive effects in service delivery may be achieved with predominantly collaborative approaches, and where data is available, there is also some evidence that such approaches might well have more sustainable outcomes (Alhassan et al., 2016; Aston, 2015; Björkman and Svensson, 2009; Donato and Mosquera, 2016; Gullo et al., 2017; Mohanan *et al.* 2020; Mwale and Msiska, 2020). Overall, then, the findings of our study cast doubt on previous arguments which suggest that “weaker” forms of citizen engagement are less effective than ‘strong enforceability (McGee and Gaventa, 2011: 11; Joshi, 2017).’

This does not, however, mean that one approach (confrontation, collaboration, or hybrid) is necessarily superior to another across different contexts. Though further study may perhaps reveal the relative advantages and disadvantages. For instance, some initiatives which pivoted to a more confrontational approach such as the HRH campaign in Uganda or the Coalition for Dignified Education (CED) in the Dominican Republic were able to achieve significant outcomes in budget commitments that might not have been achieved through a more collaborative approach during the same window of opportunity (see Larsen, 2015; Poli et al., 2019). It is also possible that learning outcomes in both India or Indonesia would not have improved to the same degree, or in the same way, had these interventions not pursued a predominantly confrontational (and punitive) approach (Duflo and Hanna, 2005; Gaduh et al., 2020). Yet, similar efforts to the HRH campaign in Peru and Guatemala did not appear to have comparable opportunities, and thus periods of pursuing more confrontational approaches were comparatively less successful, and periods or locations (Puno in Peru and San Marcos in Guatemala) in which a collaborative approach was employed were more successful (Aston, 2015, 2016; Joshi, 2017). So, rather than a generalizable statement, we suggest that some time-bound contexts may perhaps be more conducive for certain types of collaboration or confrontation than others.

---

<sup>27</sup> In one example, a council head noted ‘we help the village cadres “obtain face.” Another council member noted they were helping the village government to ‘secure face (Tsai, 2007: 112, 141).’ Thus, the potential loss of face seems a weaker motivating factor than the potential gains in moral standing secured from ‘leading the way,’ as a party secretary put it.

<sup>28</sup> There is also some evidence of similar social norms within Scheduled Caste (SC) system within Public Distribution System (PDS) in Uttar Pradesh in which communities relied on informal social networks to strengthen the accountability of public service delivery agents and ensure access to and uptake of grain, which was for SCs households (relative to non-SC households) when shopkeepers were SCs (Nagavarapu and Sekhri, 2013).

## *What contextual factors mattered most?*

Evidence of sanctioning effort was found across 17 countries. However, as few of the cases reviewed presented much information about contextual factors, patterns across a large proportion of cases were not immediately forthcoming.

We initially explored the potential of contextual variables which have been argued as enabling/disabling factors for accountability such as levels of democracy, inequality and fragility, or state capacity (Bukenya et al., 2012; Gaventa and Barrett, 2010; Gaventa and McGee, 2010; Gaventa and McGee, 2011; Joshi, 2019). This included reviewing indices of income, inequality, fragility and conflict, freedom, voice and accountability, and civic space, as well as considering potential variation by colonial history.<sup>29</sup> No strong relationship was found across cases between these features and whether civil society actors pursued a strategy which prioritized social sanctions or whether state actors attempted to impose formal sanctions. While limited civic space will likely dampen citizen voice to some degree and may also create incentives to impede formal sanctioning processes, we found no clear relationship between civic space and whether service delivery outcomes could be achieved or not via a sanction-based strategy.

Similarly, while we might also anticipate that higher capacity states would be better disposed to impose and enforce formal sanctions (see Bukenya et al., 2012; Gaventa and McGee, 2011), this was by no means a *necessary* condition. For instance, six cases included featured in the top 20 of the fragile states index at the nearest point of measurement, with four cases in contexts considered to be “unfree.” As per Gaventa and Barrett’s (2010) assessment, we found no clear relationship between the level of democracy and the level of outcomes achieved. Several cases included were in among the most unequal and poorest countries in the world, and many of these had discernibly low state capacity. Indeed, even relatively high-capacity states in the sample had significant limitations in imposing legal sanctions, with Colombia noted as an outlier exception (Herrera and Mayka, 2019; IBP and INTOSAI, 2020; Mendiburu 2020). There was no strong association with the World Bank’s voice and accountability indicators, as many contexts with the lowest rank strongly promoted formal sanctions, and some of these were associated with outcomes. Moreover, the sample includes cases from a wide variety of post-colonial and non-colonized contexts, and there was limited variation in whether mechanisms materialized. It is possible correlations might arise in a much larger sample of cases. Notwithstanding the absence of consistent trends, the research was able to identify several factors which appeared to be potentially significant, but these were only consistent across a subset of cases within the sample.

Government champions can play an important role in supporting accountability more generally (see Bukenya et al., 2012; Fox, 2014). The specific leaders that mattered varied across cases, and while their specific motivation may be unknown, these apparent champions included Members of Parliament on Uganda’s Social Services Committee<sup>30</sup> in the HRH campaign who helped increase budget allocation, the Chief Minister of the state of Andhra Pradesh who was receptive to including social audits in the MGNREGS, an undersecretary in The Philippines’ Department of Education who led a nationwide drive to ensure timely procurement and delivery of textbooks to the country’s 40,000 public schools, a Minister for Ghana’s Central Region who set specific sanitation targets around open defecation and followed up on these, and even an influential World Bank project manager in Indonesia’s KDP.<sup>31</sup> Although it is often an easy shortcut to point to supportive individual leaders, in each of these cases there was firm evidence that they played an important role.

---

<sup>29</sup> GDP per capita, Gini Coefficient, Fragile states index, Freedom House index, World Bank Voice and Accountability, CIVICUS Monitor.

<sup>30</sup> This includes outspoken MP Cernal Nebanda who played a very prominent role in the campaign (Larson, 2015), yet died in controversial circumstances only months after campaign success.

<sup>31</sup> Peru’s Minister of Health and Puno’s Ombudsman were also key supporters, but this support had no discernible connection to the role of formal or social sanctions (see Aston, 2015).

Capacity and legal authority likely make a difference to the kind of sanctions we see in the first place. For top down administrative or legal sanctions to be disbursed, evidence from Indonesia, Peru and Uganda suggest that on one hand, citizens require direct access to oversight agencies to make use of them (Buntaine and Daniels, 2020; Vargas, 2013; Woodhouse, 2005). However, without adequate staff and mandates to legally prosecute, as one representative of the Departmental Office for Integral Health Insurance (ODESIS) in Peru put it, “there is supervision without sanctions (Aston, 2015: 9; Vargas, 2013).” Yet, when citizen groups had access to a state authority who had both the mandate and capacity to impose sanctions, and assuming those authorities themselves viewed citizens’ case to be legitimate, then there appeared to be better chances of successfully disbursing sanctions, as we found in the case of the Bogotá river in Colombia, the KDP in Indonesia, a recent experiment in Uganda’s Bwindi National Park and the Accountability Can Transform Health (ACT Health) program (Bailey and Mujune, 2021; Buntaine and Daniels, 2020; Herrera and Mayka, 2019; Woodhouse, 2005).

Electoral pressures can provide windows of opportunity to impose social and formal sanctions. While there is no consistent relationship between elections and the effectiveness of sanctions, we found that electoral pressures were important in several cases (see Waddington et al., 2019).<sup>32</sup> Some initiatives took advantage of electoral windows of opportunities to name and shame governments, and these were connected to social accountability efforts themselves. In the Dominican Republic, the CED took advantage of an upcoming election (Poli et al., 2019). In Indonesia, highly competitive local elections appeared to provide a conducive environment for authorities to respond to a community protest in the Raskin (Rice for the Poor) program (Banerjee et al., 2018). In Indonesia’s KDP, audits were judged to be more effective where the village head was up for re-election and where the upcoming election was close (Olken, 2007). Electoral politics may even have had a bearing on civil servants’ receptiveness to the production of District League Tables in Ghana’s I am Aware project (Jones et al., 2019), and there was judged to be a potential relationship between electoral politics and social audits in India’s MGNREGS (Pande and Dubbudu, 2017, see also Overy, 2013 in South Africa). We may then perhaps suggest that competitive clientelist political settlements in periods of highly competitive elections provided a potentially conducive environment for naming and shaming through the media or protest to be more effective (both nationally and sub-nationally). Yet, it should be noted that many organizations will choose *not* to take advantage of electoral windows of opportunity due to the risk of being perceived as politically partisan and to mitigate the risks of being perceived as such, particularly if their preferred candidates lose.

Top-down governments may favor formal sanctions and upward accountability. On the other hand, somewhat paradoxically, yet in line with anti-corruption scholarship, dominant party settlements or more authoritarian governments appear to provide a slightly more conducive environment for imposing top-down administrative and legal sanctions on lower-level civil servants and frontline staff. This is most notable in cases from China and Uganda which lie outside our sample for their lack of citizen engagement (see Tsai et al., 2019; Ang, 2020; Peiffer et al., 2021; Tsai, 2021). Within our sample, we found evidence from Uganda, Ethiopia, and India which appears to suggest that top-heavy administrations have an appetite for imposing formal sanctions on lower levels of administration (Bailey and Mujune, 2021; Buntaine and Daniels, 2020; Larson, 2015; Pande and Dubbudu, 2017; Schjødt, 2018). However, we do not believe this association is especially strong, given that cases with an appetite for judicial sanctions were also found under competitive clientelist regimes in Nepal and Colombia (Herrera and Mayka, 2019; Schjødt, 2017).

Yet, top-down governments also appear to impede downward accountability. We found that dominant party contexts were generally less conducive to more critical or adversarial forms of citizen voice and created barriers for frontline staff escalating concerns when these contradicted top-down targets or political narratives and holding higher-level officials to account (Alemén and Burell, 2018; Aslam, 2017; Buntaine and Daniels, 2020; Pande and Dubbudu, 2017; Joshi, 2019; Tsai, 2021; Wang, 2019).

---

<sup>32</sup> See also Coehlo, 2018 for connections between strong electoral competition and administrative sanctions for contractors in São Paulo’s health system, yet without citizen engagement.

Ayliffe's (2017) evidence from social accountability efforts in Ethiopia's PSNP, for example, suggests that top-down oversight actually *deterred* frontline staff from raising concerns upwards, for fear that they would be relocated if they complained too much. So, while more authoritarian contexts may be more amenable to imposing punishment, they appear to be conducive to a suboptimal form of accountability.

A related feature to electoral pressure is the vulnerability of principals (i.e., duty bearers). While recent literature also emphasizes the importance of "countervailing power" to reduce and neutralize the power advantages of ordinarily powerful actors (Fox, 2020), we found that various state actors may be less powerful than is often assumed. Several cases reviewed demonstrated that actors took advantage of existing state efforts to decrease teachers, nurses, or civil servants' level of security, and this made them more sensitive to pressure from citizens, CSOs, managers, politicians, and donors. In India and Indonesia, evidence suggested that better-paid teachers did not perform better (Kremer et al., 2005; World Bank, 2020), and this provided a justification to introduce measures to make their lives more difficult.<sup>33</sup> The review found a greater sensitivity to pressure from formal and/or informal sanctions in a range of cases where public servants have more precarious terms of employment. Examples included a precarious contract structure in India's health sector with contract nurses (Banerjee et al., 2007) contract teachers in Kenya's education sector (Duflo et al., 2009) relocated teachers in Indonesia (Gaduh et al., 2020), short-term contractors in social protection in India (Aiyar, 2010), short-term staff in Indonesia's infrastructure sector (Woodhouse, 2005) and staff in younger and less capable districts in Ghana (Jones et al., 2019) were more sensitive to pressure from either formal or social sanctions. This suggests that sanctions are more likely to make a difference when service providers and civil servants are *already* relatively vulnerable (see also Peiffer et al., 2021 in Uganda). This also leads to an uncomfortable question about the difference sanctions really make to more powerful actors.

## Conclusions

In this paper we aimed to help bridge the gap between theory and evidence in relation to the role of formal and social sanctions in the social accountability sector. We interrogated the underlying assumptions within prevailing accountability discourse in both scholarship and among major donors which generally view sanctions as the pinnacle of accountability. We then demonstrated how these same assumptions underpin theories of change in the sector and have influenced empirical claims that sanctions are invariably necessary to achieving service delivery, and that a focus on harder forms of sanctions will likely deliver better and more sustainable results.

Overall, we found that there are significant gaps between accountability theory and empirics in evidence reviews in relation to sanctions. In reviewing the most widely cited evidence reviews, we found that there was a conspicuous gap between theoretical conceptualizations of accountability and the empirical evidence for what enhanced service delivery. We revealed that evidence reviews in the accountability sector were generally reviews of transparency and responsiveness, and these reviews rarely substantiated empirically what contribution formal or social sanctions actually made to improving service delivery.

Sanctions *can* make a difference to improved service delivery. On a close examination of 35 cases, we found that both formal and social sanctions in social accountability initiatives can make an important contribution to outcomes. We noted that a high proportion of these cases demonstrated connections between sanctions and intermediate results and a minority were also able to contribute to budget changes or downstream impacts for communities. We also revealed that these efforts were able to contribute to a wide range of changes, from increased service provider awareness to improved student test scores in language and mathematics.

---

<sup>33</sup> As Leaver and Pritchett (2019) show, the evidence base for this in the education sector is mixed.



We found that there are various contingent mechanisms of change linked to sanctions. The paper differentiated several mechanisms through which formal and social sanctions contributed to these results. In this way, we were able to make connections to behavior change more explicit. We found that these sanction-based mechanisms relied on diverse motivational impulses. Service providers and civil servants were sometimes motivated in response to punishment received, to the fear of being caught and punished by superiors, to the threat of legal punishment, or in response to a sense of shame when their misdeeds were exposed by citizens and the media. These were typically supported by service monitoring from citizen groups. We also noted that some types of actors were more common targets for sanctions than others, and we showed that stakeholders generally preferred softer sanctions for actors with whom they had close relationships and harder sanctions for those with whom they had more distant relationships.

Sanctions-based approaches were rarely sustainable and frequently triggered various negative effects. While each of the aforementioned mechanisms demonstrated the potential to influence service providers' and civil servants' behavior, we found that these results were rarely sustainable, and a high proportion of interventions which placed an emphasis on harder sanctions also contributed to a wide variety of negative effects. Our findings, therefore, challenge the assumed connection between harder forms of accountability and more significant and sustainable results (Anderson et al., 2020; Arugay, 2016; Joshi, 2010; McGee and Gaventa, 2011; Schedler, 1999; Tsai et al., 2019). In part, this may be because sanction-based approaches often fail to tackle underlying and systemic causes of service delivery problems, particularly in underfunded sectors with poorly supported staff (see Peiffer et al., 2021).

Softer approaches in our sample (and compared with cases outside the sample) were often at least as effective and sustainable and had fewer negative repercussions. Neither in evidence reviews nor in the 35 cases reviewed in this study did we find a clear connection between the intensity of sanctions and better results. In fact, we generally found the opposite; "weaker" and "softer" approaches generally had a slightly higher success rate and a significantly lower proportion of negative results. Sanction-focused approaches may indeed contribute to the achievement of some significant outcomes, but they generally have various drawbacks. And in some cases, the negative externalities from a sanctions-focused approach to accountability could be argued to outweigh the benefits (see also Peiffer et al., 2021).

Finally, we found that redeploying collaborative and confrontational tactics can be difficult. A handful of the cases reviewed were examples of hybrid strategies which employed a supposed mixture of collaborative and confrontational tactics. It has been argued that such an approach may motivate government collaboration (Anderson et al., 2020; Joshi, 2017). Yet, from the evidence reviewed in this paper, we came to the same conclusion as Fung and Wright (2003: 267, 282) that adversarial forms of engagement cannot easily be redeployed for collaborative purposes. We found that actors commonly *pivoted* from one strategy to another, and that such pivots may be more difficult to manage and may have more negative repercussions than is generally acknowledged. We thus suggest that there is likely a time and place for collaborative, confrontational, or hybrid strategies, but that none of these approaches are a panacea.

## Recommendations

A side-meeting at the Global Partnership for Social Accountability's (GPSA) annual conference in November 2019 consulted 40 practitioners, evaluators, and researchers about what they considered priorities for the sector to achieve greater impact. Participants identified three priority issues (see Guertzovich, 2020; Jacobstein, 2020). Two of these: more carefully considering civil servants' incentives and building better relationships with service providers and civil servants are highly relevant to the findings of this study and shape our recommendations for scholars and practitioners.

The social accountability field appears to have forgotten about carrots. It is time to reconsider them. In this study, we revealed how little attention has been paid to service providers and civil servants' *positive* incentives. Often, to cite Goetz and Jenkins' (2005: 192 –193) critique of one case in our sample, we found that various interventions were 'based entirely on controlling and punishing errant front-line worker[s], rather than sympathetically examining the practical constraints they face.' Many interventions reviewed focused almost entirely on removing benefits, reducing pay, disciplining, relocating, removing, or prosecuting poorly behaving staff. Scarcely any attention was afforded to identify or reward good behavior. While several cases referred to "incentive" programs, these are more accurately understood as *disincentive* programs, as the motivational impulses were not pride or fame, but fear and shame. "Carrots" should not be understood merely as the absence of punishment, but rather as the *presence* of an anticipated or real reward (see Funnel and Rogers, 2011). We, therefore, recommend greater attention to the so-called "art of bureaucraft" and studying "mission-driven bureaucrats (Joshi and McCluskey, 2017; Honig, forthcoming)," as well as further inquiry into the enabling conditions for bureaucrats and service providers not simply to refrain from behaving poorly, but also to encourage good behavior in a service delivery system. Despite the skepticism from principal-agent-inspired studies, the pay and *conditions* of service providers at the frontline remain an area which requires further consideration as indeed does state capacity to respond (see Leaver and Pritchett, 2019).

An important connected point relates to the expansive concept of teeth. Fox (2014) argues that it is better to understand the concept in broad terms. While a broad definition may be helpful to explore potential interaction effects within the concept, as we have illustrated, different stakeholders have understood the concept in contradictory ways and the concept is simply too broad to be operationally useful for accountability practitioners to plan interventions effectively. The many different components proposed need be disaggregated because they entail significantly different stakeholders, tactics, strategies, and approaches to achieve accountability and improved service delivery. Some of these may even lead to contradictory interaction effects.

Scholars, evaluators, and program teams should make better use of theory-based and participatory methods. We also noted that relatively few cases we reviewed sought to explicitly shed light on service providers' or civil servants' motivations, and rarely were the perspectives of service providers or civil servants sought to corroborate claims of the influence of sanctions. We also found that many studies failed to provide an adequate explanation of contextual features which may have influenced outcomes. As discussed by McGee and Gaventia (2011), part of this limitation is ostensibly explained by methodological biases towards Randomized Control Trials (RCTs), which generally provided inadequate attention to context and to the reasoning of service providers' or civil servants. We hope that this study has provided some greater latitude for exploration, in part, because of employing a realist approach. Reviews have become increasingly theory-based in their approach, and this should be welcomed (see e-Pact, 2016; Tsai, 2019; Waddington et al., 2019). However, reviews are necessarily constrained by their focus, and further granularity is likely beneficial for various areas of inquiry. At a bare minimum, we, therefore, recommend greater methodological pluralism in research and evaluation in the sector. And, we suggest the potential value added of theory-based methods such as realist evaluation and process tracing to unpack causal mechanisms in individual cases.

It is important to focus on supply-side actors' motivation, not just citizen voice (see Tendler, 1998). Theory-based methods will only help to shed light on the theories that researchers and evaluators aim to assess. The long-standing bias in the sector towards documenting citizen participation and voice has led to a conspicuous blind spot regarding the perspectives of service providers or civil servants. Both quantitative and qualitative studies reviewed tended to focus overwhelmingly on explanations from the demand-side stimulus rather than on the supply-side response and enquiring into why service providers and bureaucrats may change their behavior. So, we therefore strongly recommend a rebalancing of inquiry towards unpacking the drivers of motivation and responsiveness.

It is also important to further explore contextual features in individual cases. While we were unable to identify strong trends of contextual factors which enabled social and formal sanctions to play a role in

enhancing service performance, we were able to identify three conditions which we believe offer the greatest promise. These conditions were: (1) supportive leaders who played a role as champions; opening doors or accompanying civil society efforts; (2) capacity and legal authority of oversight agencies (3) competitive elections, which provided windows of opportunity for CSOs to combine political and social accountability efforts; and (4) vulnerable public servants and service providers already in relatively precarious situations. As these are preliminary findings from only a small subset of cases, we believe they offer other scholars opportunities for further exploration and empirical testing.

Finally, practitioners should carefully consider and mitigate the potential for backlash from sanctions-based approaches. Given the surprisingly high proportion of negative effects found in the cases reviewed, this is perhaps the most important recommendation of all. These risks indicate that more confrontational tactics likely come with greater risks (and costs) for a wider variety of stakeholders than is generally acknowledged, even despite the support from government oversight agencies in some instances. These risks should therefore be borne in mind by funders, INGOs, and CSOs which choose to prioritize these tactics in a particular context and at a particular moment in time. We believe there is therefore a call for more thorough political economy analysis at project or campaign inception. Indeed, as pathways to change are highly contingent, we also recommend the monitoring of conflict dynamics during initiatives or campaigns to limit the potential harms which may arise from more adversarial approaches.

## Bibliography

- Aceron, J. 2016. "Mobilizing Citizens for Transparency and Accountability: The Textbook Count Experience," in J. Fox and J. Aceron, with A. Guillán Montero 'Doing Accountability Differently: A Proposal for the Vertical Integration of Civil Society Monitoring and Advocacy', *U4 Issue Paper No. 4*, August.
- Aceron, J. 2018. "Going Vertical: Citizen-led Reform Campaigns in the Philippines," Quezon City: Government Watch (G-Watch).
- Ackerman, J. 2005. "Social Accountability in the Public Sector A Conceptual Discussion," Social Development Papers, Participation and Civic Engagement Paper No. 82, March 2005.
- Aiyar, Y. 2010. "Invited Spaces, Invited Participation: Effects of Greater Participation on Accountability in Service Delivery." *India Review*, Vol. 9, No. 2, pp. 204-229.
- Agarwal, S., Heltberg, R. and Diachok, M. 2009. "Scaling up Social Accountability in World Bank Operations," World Bank.
- Alhassan, R. K., Nketiah-Amponsah, E., Spieker, N., Arhinful, D. K., and Rinke de wit, T. F. 2016. "Assessing the Impact of Community Engagement Interventions on Health Worker Motivation and Experiences with Clients in Primary Health Facilities in Ghana: A Randomized Cluster Trial," *Plos One*, Vol. 11, No. 7.
- Almén, O., and Burell, M. 2018. "Social Accountability as Social Movement Outcome: Protests in a Chinese City," *Social Movement Studies*, Vol. 17, No. 6, pp. 716-735.
- 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, Vol. 2020, No. 536.
- Ang, Y. Y. 2020. *China's Gilded Age: The Paradox of Economic Boom and Vast Corruption*, London: Cambridge 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," Faculty Research Working Paper Series, Harvard Kennedy School.
- Arroyo, D. and Sirker, K. 2005. "Stocktaking of Social Accountability Initiatives in the Asia and Pacific Region," World Bank, Washington, D.C.
- Arugay, A. 2016. "Sanctions, Rewards and Learning: Enforcing Democratic Accountability in the Delivery of Health, Education, and Water, Sanitation and Hygiene," International IDEA Discussion Paper 16/2016.
- Aslam, G. 2018. "Social Accountability in the Delivery of Social Protection: South Africa Case Study," Development Pathways: London.
- Aston, T. 2015. "Citizen Monitoring to Defend Maternal Health Rights in Peru," CARE International.
- Aston, T. 2018. "What has CARE learned from political economy analysis?" Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP).
- Aston, T. 2019. "CARE International Advocacy and Influencing: A Review of Pathways to Success," CARE International.
- Aston, T. 2020. "W(h)ither Sanctions," LinkedIn.
- Aston, T. 2021. "What Actually is Social Accountability," Medium.
- Ayliffe, T. 2017. "Social Accountability in Social Protection: Ethiopia Case Study," Development Pathways: London.

- Ayliffe, T., Aslam, G., and Schjødt, R. 2017. "Social Accountability in the Delivery of Social Protection: Literature Review," Development Pathways: London.
- Bailey, A. and Mujune, V. 2021. "Bottom-up Accountability in Uganda: Learning from People-centered, Multi-level Health Advocacy Campaigns," Working Paper Number 8, Accountability Research Center.
- 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*, Vol. 6, pp. 487–500.
- Banerjee, A., Hanna, R., Kyle, J., Olken, B. A., and Sumarto, S. 2018. "Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia," *Journal of Political Economy*, Vol. 126, No. 2, pp. 451–491.
- Batzin, B, Culum, P. and Fischer-Mackey. J. 2020. "Defending the Right to Health in Guatemala: Reflections of Two Indigenous Women on the Frontlines," Accountability Note 8, Accountability Research Center.
- Beach, D. and Pedersen, R.B. 2019. *Process-Tracing Methods: Foundations and Guidelines*, Ann Arbor MI: University of Michigan Press.
- Beasley, E. and Huillery, E. 2013. "Empowering Parents in School: What They Can (not) Do," "Sciences Po Economics Discussion Papers, 2013, Vol. 3.
- Behn, R. D. 2001. *Rethinking Democratic Accountability*. Washington, D.C.: Brookings Institution 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*, Vol. 124, pp. 735–769.
- Björkman Nyqvist, M., deWalque, D., and Svensson, J. 2014. "Information is Power: Experimental Evidence on the Long-Run Impact of Community Based Monitoring," World Bank Group.
- Boas. T. Hidalgo, D. 2019. "Electoral Incentives to Combat Mosquito-borne Illnesses: Experimental Evidence from Brazil," *World Development*, Vol. 113, pp. 89–99.
- Bovens, M. 2005. "Public Accountability: A Framework for the Analysis and Assessment of Accountability Arrangements in the Public Domain," Draft made for CONNEX, Research Group 2: Democracy and Accountability in the EU, in E. Ferlie, L. Lynne & C. Pollitt (eds.), *The Oxford Handbook of Public Management*, Oxford: Oxford University Press.
- Bovens, M. Goodin, R. and Schillemans, T. 2014. *The Oxford Handbook of Public Accountability*, Oxford: Oxford University Press.
- 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, Washington D.C.: Population Council.
- Bradley, J. and Igras, S. 2005. "Improving the quality of child health services: participatory action by providers," *International Journal for Quality in Health Care*; Vol. 17, No. 5, pp. 391–399.
- Brook, S. Holland, J. and Macuane, J. 2017. Mozambique Citizen Engagement Programme (CEP): Final Evaluation Report, Oxford: Oxford Policy Management.
- Buntaine M. and Daniels, B. 2020. "Combining bottom-up Monitoring and Top-down Accountability: A Field Experiment on Managing Corruption in Uganda," *Research and Politics*, July-September 2020, pp. 1–9.
- Bukenya, B. Hickey, S, and King, S. 2012. "Understanding the Role of Context in Shaping Social Accountability Interventions: Towards an Evidence-based Approach," World Bank.

- Caldeira, E. 2012. “Yardstick Competition in a Federation: Theory and Evidence from China,” *China Economic Review*, Vol. 23, No. 4, pp. 878-897.
- Carr-Hill R, Rolleston C, Schendel R. 2016. “The Effects of School-based Decision Making on Educational Outcomes in Low- and Middle-income Contexts: A Systematic Review,” *Campbell Systematic Reviews*, 2016:9.
- Centre for the Study of Equity and Governance in Health Systems – CEGSS. 2015. “Plataforma de Denuncias Sobre Servicios Públicos de Salud: Boletín Informativo,” CEGSS.
- Chen, Y. Lee, P. and Lu, Y. 2015. “Accountability, Career Incentives, and Pollution: The Case of Two Control Zones in China,” Netherlands: SSRN.
- Coelho, V. S. 2018. “Reducing Health Inequalities in Brazil’s Universal Health-Care System: Accountability Politics in São Paulo,” *IDS Bulletin*, Vol. 49, No. 2, pp. 109- 128.
- Cunil-Grau, N. 2008. “Contraloría Social y Derechos Sociales El Desafío de la Integralidad,” *Gestión y Política Pública*, Vol. 18, No. 1, pp. 3 – 37.
- Department for International Development – DFID. 2014. “Assessing the Strength of Evidence: How to Note, Department for International Development.
- Donato, K. and Mosquiera, A. 2016. “Power to the People? A Replication Study of Community-based Monitoring in Uganda,” Replication Paper 11, Washington, D.C. International Initiative for Impact Evaluation.
- Duflo, E. and Hanna, R. 2005. “Monitoring Works: Getting Teachers to Come to School,” NBER Working Paper No. 11880.
- Duflo, E., Dupas, P. and Kremer, M. 2009. “Additional Resources Versus Organizational Changes in Education: Experimental Evidence from Kenya, Unpublished manuscript. Abdul Latif Jameel Poverty Action Lab (JPAL), Cambridge, Mass.: Massachusetts Institute of Technology.
- Duflo E, Hanna R, and Ryan S. 2010. “Incentives Work: Getting Teachers to Come to School,” Abdul Latif Jameel Poverty Action Lab (JPAL), Cambridge, Mass.: Massachusetts Institute of Technology.
- Edwards, D. Lopez Franco, E. and Shutt, C. 2020. “Open Up Contracting: End-Term Evaluation of the Citizen Agency Consortium Open Up Contracting Program,” The TAP Room.
- E-Pact Consortium. (2016). “Empowerment and Accountability Annual Technical Report 2016: What Works for Social Accountability, Macro Evaluation of DFID’s Policy Frame for Empowerment and Accountability,” Oxford: e-PACT.
- Feinglass, E. Gomes, N. and Maru, V. 2016. “Transforming Policy into Justice: The Role of Health Advocates in Mozambique,” *Health and Human Rights Journal*, Vol. 18, No. 2, pp. 233-246.
- Ferraz, C. and Finan, F. 2008. “Exposing Corrupt Politicians: The Effects of Brazil’s Publicly Released Audits on Electoral Outcomes,” *The Quarterly Journal of Economics*, May 2008.
- Feruglio, F. 2017. “Do More Empowered Citizens Make More Accountable States? Power and Legitimacy in Legal Empowerment Initiatives,” Making All Voices Count Research Report, Brighton: Institute for Development Studies.
- Fox, J. 2007. “The Uncertain Relationship between Transparency and Accountability,” *Development in Practice*, Vol. 17, No. 4, pp. 663 - 671.
- Fox, J. 2014. “Social Accountability: What does the evidence really say?” GPSA Working Paper No. 1, Washington, DC: International Bank for Reconstruction and Development.
- Fox, J. 2016. Scaling Accountability through Vertically Integrated Civil Society Policy Monitoring and Advocacy, Working Paper, Making All Voices Count.

- Fox, J. 2018. “The Political Construction of Accountability Keywords,” *IDS Bulletin*, Vol. 49, No. 2, pp. 65 – 81.
- Fox, J. 2020. “Contested Terrain: International Development Projects and Countervailing Power for the Excluded,” *World Development*, Vol. 132 (2020).
- Flores, W. 2018. “How Can Evidence Bolster Citizen Action? Learning and Adapting for Accountable Public Health in Guatemala,” *Accountability Note 2*, Accountability Research Center.
- Flores, W. and Hernández, A. 2018. “Health Accountability for Indigenous Populations: Confronting Power through Adaptive Action Cycles,” *IDS Bulletin*, Vol. 49, No. 2, pp. 19 – 34.
- Fung, A. and Wright, E.O. 2003. *Deepening Democracy: Institutional Innovations in Empowered Participatory Governance*, The Real Utopias Project 4, London: Verso.
- Fung, A. 2020. “Afterword: Does Deliberative Democracy Have a Role in Our Time of Political Crisis?” *Journal of Deliberative Democracy*, Vol. 16, No. 1, pp. 75–76.
- Funnell, S. and Rogers, P. 2011. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*, San Francisco: John Wiley & Sons, Inc.
- 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 for Development Studies.
- Gaventa, J., and McGee, R. 2010. *Citizen Action and National Policy Reform: Making Change Happen*, London: Zed Books.
- Goetz, A.M. and Jenkins, R. 2005. “Reinventing Accountability: Making Democracy Work for Human Development,” London: Palgrave Macmillan.
- 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.
- Grossman, G., Jonathan, R., Tausanovich, Z., and Han, A. 2017. “Can Text Messages Improve Local Governance? An Impact Evaluation of The U-Bridge Programme In Uganda,” United States Agency for International Development.
- Gruenberg, C., and Pereyra Iraola, V. 2008. “Sistemas de Rendición de Cuentas: De la Teoría Promisoria a la Práctica Concreta: Los Casos de México y Argentina,” in *Candados y Contrapesos: La Protección de los Programas, Políticas y Derechos Sociales en México y América Latina*, Ed. David Gómez Álvarez, 169–94. Tlaquepaque, Mexico: Instituto Tecnológico y de Estudios Superiores de Occidente.
- Guerzovich, F. 2020. “Conquering our Blind Spots: Glamorizing the Everyday Politics of Accountability (Part 1),” Global Partnership for Social Accountability.
- Guerzovich, 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.
- Guerzovich, F. and Schommer, P. 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.
- Gullo S, Galavotti C, Sebert Kuhlmann A, Msiska T, Hastings P, Marti, C. 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*, Vol. 12, No. 2.
- 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*, pp. 1 – 12.

- Hagen-Zanker, J. and Mallett, R. 2013. “How to do a Rigorous, Evidence-focused Literature Review in International Development: A Guidance Note,” Working Paper, Overseas Development Institute.
- Halloran, B. 2015. “Strengthening Accountability Ecosystems: A Discussion Paper,” Think Piece Series, Transparency and Accountability Initiative.
- Hanna, R., Bishop, S., Nadel, S., Scheffler, G, Durlacher, K. 2011. “The Effectiveness of Anti-corruption Policy: What has Worked, what Hasn’t, and what we don’t Know – A Systematic Review,” Technical Report. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Hepworth, N.D., Brown, B.D. and Brewer, T. 2020. “Accountability and Advocacy Interventions in the Water Sector: A Review of Global Evidence, Part 1. Overview and Summary Results. Water Witness International, Edinburgh, UK.
- 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* (2020) 20:416.
- Hernández, A. Ruano, A. Hurtig, A. Goicolea, I. San Sebastián, M. and Flores, W. 2019. “Pathways to Accountability in Rural Guatemala: A Qualitative Comparative Analysis of Citizen-led Initiatives for the Right to Health of Indigenous Populations, *World Development*, Vol. 113, pp. 392 – 402.
- Herrera, V. and Mayka, L. 2019. “How Do Legal Strategies Advance Social Accountability? Evidence from Colombia,” *The Journal of Development Studies*, Vol. 56, No. 8, pp. 1437-1454.
- Hevia, F. J. and Vergara-Lope, S. 2019. “Educational Accountability or Social Accountability in Education? Similarities, Tensions, and Differences.” Accountability Working Paper 5, Accountability Research Center.
- Honig, D. forthcoming. *Mission Driven Bureaucrats*, 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.” CGD Working Paper 510, Washington, DC: Center for Global Development.
- International Budget Partnership and International Organization of Supreme Audit Institutions 2020. “All Hands on Deck: Harnessing Accountability through External Public Audits: An Assessment of National Oversight Systems,” IDI and IBP.
- Jacobstein D. 2020. “What is the Work?” USAID Learning Lab.
- Jones, E. Amidu, I-T, and Sefa-Nyarko, C. 2019. Formative Evaluation of the ‘I Am Aware’ Social Accountability Project in Ghana, Oxford Policy Management (OPM).
- Joshi, A. 2010. “Annex 1: Review of Impact and Effectiveness of Transparency and Accountability Initiatives,” Transparency and Accountability Initiative.
- Joshi. A. 2017. “Legal Empowerment and Social Accountability: Complementary Strategies Toward Rights-based Development in Health?” *World Development*, Vol. 99, pp. 160–172.
- Joshi, A. 2019. “Introduction: Accountability Amidst Fragility, Conflict, and Violence: Learning from Recent Cases,” *IDS Bulletin*, Vol. 50, pp. 1-17.
- Joshi, A., and Houtzager, P. 2012. “Widgets or Watchdogs?” *Public Management Review*, Vol. 14, No. 2, pp. 145–162.
- Joshi, A. and McCluskey, R. 2017. The Art of “Bureaucraft”: Why and How Bureaucrats Respond to Citizen Voice, Making All Voices Count Research Briefing, Brighton: Institute for Development Studies.
- Kirk, T. 2017. “A Strategic Approach to Social Accountability in Pakistan,” IDS Working Paper 497, Brighton: Institute for Development Studies.



- Klugman, J. 2013. “Case Study Report: Promoting Citizens’ Engagement in Basic Education,” Australian Community Development and Civil Society Strengthening Scheme (ACCESS) Phase II, Canberra, AusAID.
- Kosack, S. and Fung, A. 2014. “Does Transparency Improve Governance?” *Annual Review of Political Science*, Vol. 17, pp. 65-87.
- Kosec K. and Wantchekon, L. 2020. “Can Information Improve Rural Governance and Service delivery?” *World Development*, Vol. 125.
- Kremer, M., Chaudhury, N., Rogers, F., Muralidharan, K. and Hammer, J. 2005. “Teacher Absence in India: A Snapshot,” *Journal of the European Economic Association*, Vol. 3, pp. 658–667.
- Larsen, J. 2015. “Uganda: Winning Human Resources for Health. Case Study,” Washington, DC: International Budget Partnership.
- Leaver, C. and Pritchett, L. 2019. “Should Pay for Performance be used for teachers? (With a plea to pause before you answer),” available at: <https://riseprogramme.org/blog/pay-for-performance-plea-pause> (accessed 10/08/22).
- Le Billion, P. Lujala, P. and Rustad, S. 2020. “A Theory of Change for the Extractive Industries Transparency Initiative: Designing Resource Governance Pathways to Improve Developmental Outcomes,” U4 Special Edition.
- Lodenstein E, Ingemann C, Molenaar JM, Dieleman M, Broerse J.E.W. 2018. “Informal Social Accountability in Maternal Health Service Delivery: A Study in Northern Malawi,” *Plos One*, Vol. 13, No. 4.
- Majeed, R. 2011. “Promoting Accountability, Monitoring Services: Textbook Procurement and Delivery, The Philippines, 2002-2005,” *Innovations for Successful Societies*, Princeton University.
- Malena, C. Forster, R. and Singh, J. 2004. “Social Accountability: An Introduction to the Concept and Emerging Practice,” *Social Development Paper 76*, Washington DC: World Bank.
- Mansbridge, J. 2014. “A Contingency Theory of Accountability,” in *The Oxford Handbook of Public Accountability*, Eds. Bovens, Mark, Robert E. Goodin, Thomas Schillemans, London: Oxford University Press.
- McGee, R., and Gaventa, J. 2011. *Shifting Power? Assessing the Impact of Transparency and Accountability Initiatives*, IDS Working Paper. Brighton: Institute of Development Studies.
- McNeil, M. and Malena, C. 2010. *Demanding good governance: lessons from social accountability initiatives in Africa*. Edited by Mary McNeil and Carmen Malena, World Bank.
- 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.
- Ministry of Federal Affairs and Local Development. 2016. “LGCDP Newsletter, Local Governance and Community Development Programme-II, Jan-Jun 2016/Asadh 2073,” Vol. 1, Iss. 1, Government of Nepal, Ministry of Federal Affairs and Local Development.
- Molina, E. 2014. “Can Bottom-up Institutional Reform Improve Service Delivery?,” Washington, DC: Inter-American Development Bank.
- Molina, E. Carella, L. Pacheco, A.; Cruces, G; Gasparini, L. 2017. “Community Monitoring Interventions to Curb Corruption and Increase Access and Quality in Service Delivery,” *Journal of Development Effectiveness*, Vol. 9, Issue 4.
- Mohanam, M. and Rajan, V. S. and 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 No. 295*, February 2020.

- Moore, M. and G. Teskey 2006. “The CAR Framework: Capability, Accountability, Responsiveness: What Do These Terms Mean, Individually and Collectively? A Discussion Note for DFID Governance and Conflict Advisers,” DFID Governance Retreat, 14-17 November 2006.
- Muligan, R. 2000. “Accountability: An Ever-expanding Concept?” *Public Administration*, Vol. 78, No. 3, pp. 555–573.
- Mwale, P. and Msiska, T 2020. “Government-led Community Score Card© for Family Planning Services,” CARE International.
- Nagavarapu, S. and Sekhri, S. 2013. “Informal Monitoring Mechanisms in Public Service Delivery: Evidence from the Public Distribution System in India,” *Journal of Development Economics*, Vol. 121, pp. 63–78.
- 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*, Vol. 3, No. 4).
- Olken, B. 2007. “Monitoring Corruption: Evidence from a Field Experiment in Indonesia,” *Journal of Political Economy*, Vol. 115, No. 2, pp. 200–249.
- Ortiz, I., Burke, S. Berrada, M, Saenz Cortés, H. 2022. *World Protests: A Study of Key Protest Issues in the 21<sup>st</sup> Century*, Friedrich-Ebert-Stiftung New York; Initiative for Policy Dialogue/Global Social Justice.
- Overy, N. 2013. “The Social Justice Coalition and Access to Basic Sanitation in Informal Settlements in Cape Town, South Africa,” From Analysis to Impact Partnership Initiative Case Study Series, International Budget Partnership.
- Pande, S. and Dubbudu. R. 2017. “Citizen Oversight and India's Right to Work Program: What Do the Social Auditors Say?” Accountability Working Paper 1, Accountability Research Center.
- Pandey, P., Goyal, S. and Sundararman, V. 2009. “Community Participation in Public Schools: Impact of Information Campaigns in Three Indian States,” *Education Economics*, Vol. 17 No. 3, pp. 355-375.
- Paul, S. 1992. “Accountability in Public Services: Exit, Voice, and Control,” *World Development*, Vol. 20, No. 7, pp. 1047–1060.
- Pawson R. and Tilley N. 1997. *Realistic Evaluation*. London: Thousand Oaks and Sage.
- Peiffer, C. Armytage, R. Marquette, H. and Gumisiriza, P. 2021. “Lessons from Reducing Bribery in Uganda’s Health Sector,” *Development Policy Review*, pre-print.
- Peruzzotti, E. and Smulovitz, C. 2006. “Social Accountability: An Introduction,” In E. Peruzzotti & C. Smulovitz (Eds), *Enforcing the Rule of Law: Social Accountability in the New Latin American Democracies*, Pittsburgh: University of Pittsburgh Press.
- Poli, M. Guerzovich, F. and Fokkelman, E. 2019. “Aligning Actors for Learning: The Evolution of Social Accountability in the Dominican Republic’s Education Sector,” GPSA Note 14, Global Partnership for Social Accountability.
- Raffler, P. 2019. “Does Political Oversight of the Bureaucracy Increase Accountability? Field Experimental Evidence from an Electoral Autocracy.”
- Raffler, P., Posner, D. and Parkerson, D. 2019. “The Weakness of Bottom-Up Accountability: Experimental Evidence from the Ugandan Health Sector.”
- Ravallion, M., van deWalle, D., Dutta, P., and Murgai, R. 2013. ‘Testing Information Constraints on India's Largest Antipoverty Programme/Testing Information Constraints on India's Largest Antipoverty Program,’ Washington, D.C.: The World Bank.

- Reinikka R., Svensson J. 2003. “The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture,” World Bank Policy Research Working Paper Series 3239, Washington, DC: World Bank.
- Republic of Uganda. 2013. “National Budget Framework Paper FY 2013/14 – FY 2017/18,” Republic of Uganda.
- Ringold, D., Holla, A. Koziol, M. and Srinivasan, S. 2012. “Citizens and Service Delivery: Assessing the Use of Social Accountability Approaches in Human Development,” *Direction in Development: Human Development*. Washington DC: The World Bank.
- Rocha Menocal, A, and Sharma, B. 2008. “Joint Evaluation of Citizen’s Voice and Accountability Report,” London, UK: Department for International Development.
- 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* (2020) 20:1084.
- Schedler, A. 1999. “Conceptualizing Accountability,” in Andreas Schedler, Larry Diamond, and Marc F. Plattner, Eds., *The Self--Restraining State: Power and Accountability in New Democracies*, Boulder, CO: Lynne Rienner.
- Schjødt, R. 2017. “Social Accountability in Social Protection: Nepal Case Study,” *Development Pathways*: London.
- Schjødt, R. 2018. “Social Accountability in the Delivery of Social Protection: India Case Study,” *Development Pathways*: London.
- Squires F, Martin Hilber A, Cordero JP, Boydell V, Portela A, Lewis Sabin M. 2020. “Social Accountability for Reproductive, Maternal, Newborn, Child and Adolescent Health: A Review of Reviews,” *Plos One*, Vol. 15, No.10.
- Staplehurst R. and O’Brien, M. 2006. “Accountability in Governance,” Briefing Note, World Bank.
- Strom, K., 2000. “Delegation and Accountability in Parliamentary Democracies’, *European Journal of Political Research*, Vol. 37, pp. 261-289.
- Strom, K., 2003. “Parliamentary Democracy and Delegation,” in: K. Strom et al.(eds.), *Delegation and Accountability in Parliamentary Democracies*, Oxford: Oxford University Press.
- Taylor, G. and Khan, S. 2017. “Empowerment, Voice and Accountability for Better Health and Nutrition (EVA-BHN) – Independent Assessment,” Oxford: HEART Consortium.
- 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*, Suppl 2: 367.
- Tendler, J. 1998. *Good Government in the Tropics*, Baltimore: The Johns Hopkins University Press/
- Tsai, L. 2007. *Accountability without Democracy: Solidary Groups and Public Goods Provision in Rural China*, Cambridge, MA: Cambridge University Press.
- Tsai, L. 2021. *When People Want Punishment: Retributive Justice and the Puzzle of Authoritarian Popularity*. Cambridge, MA: Cambridge University Press.
- Tsai, L. Morse, B. Toral, G and Lipovsek, V. 2019. “Information and Accountability: Evidence Synthesis of Within Government and Citizen-Government Accountability Pathways,” MIT Governance Lab.
- United Nations Educational, Scientific and Cultural Organisation – UNESCO. 2017. “Accountability in Education: Meeting our Commitments,” United Nations Educational, Scientific and Cultural Organisation.
- 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, John D, Stevenson J. 2019. "Citizen Engagement in Public Services in Low- and Middle-income Countries: A Mixed Methods Systematic Review of Participation, Inclusion, Transparency and Accountability (PITA) initiatives," *Campbell Systematic Reviews*. 2019;15: e1025.
- Walker, D. 2017. "Leveraging Communities' Capabilities to Increase Accountability for Health Rights: The Case of Citizen Voice and Action," *Journal of Human Development and Capabilities*, Vol. 19, No. 2, pp. 181-197.
- Wang, E. 2019. "Frightened Mandarins: The Adverse Effects of Fighting Corruption on Local Bureaucracy, SSRN 3314508.
- Wein, T. 2020. "Intervening Intensely: How to Estimate Dosage in Accountability Interventions," *Medium*.
- 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.
- Woodhouse, A. 2005. "Village corruption in Indonesia: Fighting corruption in the World Bank's Kecamatan Development Program Indonesia," *Social Development Paper*; No. 6. Washington, DC: The World Bank.
- World Bank. 2003. "World Development Report 2004: Making Services Work for Poor People," Washington, DC: World Bank.
- World Bank 2020. "Community Participation and Teacher Accountability: Improving Learning Outcomes in Remote Areas of Indonesia," World Bank.
- Zeitlin, A., Bategeka, L., Guloba, M., Kasirye, I. and Mugisha, F. 2011. "Management and Motivation in Ugandan Primary Schools: Impact Evaluation Final Report," Centre for the Study of African Economies, University of Oxford.
- Zheng, S. Kahn, M. Sun, W. and Luo, D. 2013. "Incentivizing China's Urban Mayors to Mitigate Pollution Externalities: The Role of the Central Government and Public Environmentalism," NBER Working Paper No. 18872.