

Transparency for Development:  
Intervention Design and Evaluation Approach

Discussion Paper for E-GAP Meeting

London, June 3, 2013

This draft: May 27, 2013

This discussion paper outlines a recently started mixed method research program to assess whether, why, and in what contexts transparency and accountability (T/A) interventions improve health outcomes. The project is intended to advance the state of knowledge about the impact of T/A interventions on service delivery in several ways. First, we will develop a new T/A intervention, based on an initial theory about how T/A affects service delivery in different contexts as well as the range of existing work in this field, that is designed to be flexible enough to work across multiple contexts. Development of this intervention will leverage the local knowledge of embedded and capable local partners through an intensive co-design process. Second, we will conduct mixed-methods evaluations of this intervention in multiple sites: our research plan combines an initial phase in which we will evaluate its impact in two carefully selected sites, Tanzania and Indonesia, with both a randomized controlled trial and an integrated and extensive qualitative evaluation to understand mechanisms and the role of context; a cost-effectiveness analysis; and a second phase, which explores the generalizability of both our first phase results and the theoretical implications of them for mechanisms and the role of context. Third, these findings will provide the basis for development of a more nuanced and empirically grounded theory of the impact of T/A—both the different mechanisms through which T/A interventions seek to affect development outcomes, such as collaborative problem solving or community pressure, and the conditions under which these mechanisms should be expected to improve service delivery. All of these elements will be used to produce a range of new tools for practitioners of T/A as well as sectoral health experts assessing whether to add T/A interventions instead of or alongside more traditional health interventions.

## I. INTRODUCTION: CAN TRANSPARENCY IMPROVE HEALTH CARE?

Recent decades have seen vast expansions of health services across much of the developing world. But the quality of these services is often lacking, limiting their ultimate impact on the world's poor. One reason is poor governance: lack of positive incentives for frontline service providers to perform well; leakage and diversion of materials and funds; poor management; ineffective operations; and low skill levels. “Transparency and accountability” (T/A) is increasingly seen as one solution to some of these problems. The intuitive idea is that disclosure of information about the organization, operations, or provision of public services may cause providers and policymakers to improve the quality of those services and empower communities to hold these public agents to account, potentially improving the quality of public services and of governance more broadly.

This approach appears to hold substantial promise. The celebrated intervention in Björkman and Svensson (2009), for example, seems to have improved the delivery of health care in Uganda to such an extent that it reduced under-5 mortality in treatment communities by 33 percent in just a year, as well as improving treatment practices, increasing immunization rates and utilization of general outpatient services, and reducing waiting times and practitioner absenteeism—all without providing additional health care resources. Yet a decade of research into the impact of T/A interventions, including a large number of randomized controlled trials, has failed to produce any consensus on whether these interventions generally have beneficial impacts, much less on the extent of these benefits; nor has research provided an empirically verified account of what sorts of

T/A strategies are best suited to various conditions and contexts. Some studies show positive effects of T/A interventions on community empowerment and health and education outcomes, while others show no impact.<sup>1</sup> And while many proponents are hopeful that T/A interventions can help to empower the poor, some critics view these kinds of interventions as marginal and largely irrelevant to the deeper problems of power inequity, institutional failure, or social conflict that often foster slow and uneven development.<sup>2</sup> One consequence of this evidence gap is that reformers, policymakers, donors, and other practitioners lack clear guidance on what sorts of T/A interventions should be pursued—or, indeed, whether any T/A interventions can ultimately improve service delivery and health outcomes for poor populations.

In this paper we outline a two-phase, mixed method research program to assess whether, why, and in what contexts transparency and accountability interventions improve health care and health outcomes. Section II outlines a preliminary theory on how transparency and accountability interventions might affect health outcomes in different contexts. Section III describes a process of intensive co-design with local CSO partners of a T/A intervention with the flexibility to work across different contexts. Section IV details our research plan for both evaluating the causal effects of the intervention and gaining a deep and practically-applicable understanding of the roots of these effects and their likely generalizability. The research program combines an initial phase in which we will evaluate our intervention in two carefully selected sites, Tanzania

---

<sup>1</sup> See Andrabi, Das, and Khwaja (2009), Hastings and Weinstein (2008), Pandey et al. (2007), Pandey, Goyal, and Sundararaman (2009), Keefer and Khemani (2011), Björkman and Svensson (2009), Pradhan et al. (2011b), Duflo, Hanna, and Ryan (2012), Banerjee et al. (2010), Glewwe and Maïga (2011), Muralidharan and Sundararaman (2010), Olken (2007), Reinikka and Svensson (2005, 2011), Banerjee, Duflo, and Glennerster (2008).

<sup>2</sup> McGee and Gaventa (2011), Joshi and Houtzager (2012), Fox (2007), Fung, Graham, and Weil (2007).

and Indonesia, with both a randomized controlled trial and an integrated and extensive qualitative evaluation to understand mechanisms and the role of context; a cost-effectiveness analysis; and a second phase, which explores the generalizability of both our first phase results and the theoretical implications of them for mechanisms and the role of context. Section V concludes.

## II. THE ROLE OF CONTEXT

Transparency and accountability interventions are increasingly wrestling with context, which appears to account for at least some of the variation in the success of accountability interventions (e.g., Mansuri and Rao 2013, McGee and Gaventa 2011). But which contextual factors matter? In the case of transparency and accountability interventions, the important contextual factors are likely to include those that intervene in the series of principal-agent relationships that are involved in the delivery of health services.<sup>3</sup> Severe asymmetries of information and expertise between principal and agent often increase the difficulty of ensuring quality service provision, particularly in health care; two oft-cited consequences are widespread absenteeism (Chaudhury et al. 2005) and low provider effort (Das, Hammer, and Leonard 2008).

The “transparency” piece of T/A interventions is designed to provide citizens with information that is useful in reducing these kinds of problems. In addition, by involving

---

<sup>3</sup> Citizens delegate responsibility for providing public health care to politicians, who in turn delegate to ministries, who delegate to bureaucrats and finally to front-line service providers. Even where the first of these relationships works well—even where politicians are highly motivated to respond to citizen demands for quality public health care, for example because it is important to their re-election or to powerful interest groups in the society (Ross 2006, Beall and Ngonyama 2009, Kosack 2012, McGuire 2010, Zweifel and Navia 2000)—politicians may not be able to similarly motivate and monitor agents further down the accountability chain.

citizens, not (or not exclusively) bureaucrats and politicians, T/A interventions differ from other methods for improving public services in that they often they also seek a more lasting dynamic of mutually reinforcing improvements in public-sector accountability and citizen empowerment (Narayan and Petesch 2007).<sup>4</sup> The effectiveness of public services like health care depends at least in part on their being accountable to the needs of citizens; but accountability to citizen demands in public service delivery is generally complicated by the lack an “exit” option. Without this option, citizens must rely on voice,<sup>5</sup> the effectiveness of which depends on citizens—perhaps individually but more often collectively—being willing and capable of expressing their demands and acting if they are not met.<sup>6</sup> By involving citizens directly in monitoring government performance, T/A techniques can allow citizens to gain first-hand information about problems such as lack of provider effort, resources, or know-how (Ackerman 2004, Joshi 2008).

What sort of information is salient to citizens in evaluating the quality of providers and seeking improvements? Is that information useful to individual citizens, or is it only useful if citizens are able to mobilize and act collectively? And *how* is that information useful? Does it help citizens choose different providers? Does it help them work with providers to diagnose and alleviate problems in service quality? Does it help them work with those further up the accountability chain—bureaucrats or politicians who may share an interest in improving service delivery? Or does it help them to place

---

<sup>4</sup> The gradual process to which T/A intervention seek to contribute is akin to the “step changes” rubric recently developed by DFID (e.g., DFID 2011) in which poor citizens gradually exercise greater choice and control over their lives and capacities and gradually hold politicians and governments to greater account for designing and implementing policies that help citizens with that process. The contribution of the T/A interventions that we envision in this project is primarily bottom-up; however, as we describe in detail later, they also seek to engage with existing top-down or horizontal accountability dynamics or, through bottom-up pressure, to induce such dynamics.

<sup>5</sup> Albert Hirschman (1970) famously described the difficulties with relying on voice without the exit option.

<sup>6</sup> This sort of empowerment—of citizens’ voice over service delivery—is most often associated political empowerment, as opposed to economic and social empowerment (Eyben, Kabeer, and Cornwall).

pressure on either frontline providers or bureaucrats and politicians?

It is in trying to answer questions like these that the particular role for context becomes clear. A core research objective of our project is to produce a theory that delineates answers to these questions depending on the political economy context. Here we outline an early-stage theory—currently under development—from which we derive several general hypotheses for the mechanisms by which transparency may lead to improved health care in different political economy contexts. In the T4D project, this preliminary theory will be tested, inductively augmented and amended, and then tested again—a process we describe below in Section IV.

The theory builds upon the already well-known framework of long- and short-routes to accountability described in the 2004 *World Development Report* (World Bank 2004) as well as the current literature on the role and effects of transparency and accountability interventions and additional insights from broader literatures in economics and political science. The result is a preliminary rubric of “five worlds” of service delivery: varied contexts in which transparency and accountability interventions should be expected to affect service delivery, citizen empowerment and action, and development outcomes differently. In each of these worlds, the role of information and transparency is different, and each therefore calls for different kinds of reform strategies.

The differences stem from a set of choices inherent in transparency and accountability interventions. The most familiar of these is whether the information the intervention provides should help citizens and their allies to target public officials at the local level or at higher levels of government (such as provincial or national levels): the frontline providers of government goods and services or the policymakers and politicians

who oversee them. This choice is at the root of the “long” and “short” routes of accountability in the World Bank’s 2004 *World Development Report* on “Making Service Work for Poor People”—one of the most influential frameworks in the T/A arena. In the “short route,” citizens engage directly with those who provide public services, to press for the improvement of those services. In the “long route,” citizens use their political power—voting or advocacy—to press policymakers and politicians to use their positions of oversight to make government services more efficient and effective.

Whether to pursue the long or short route is, however, only one of the important choices in designing effective Transparency and Accountability programs. An equally important choice is whether the engagement should be collaborative or confrontational. This question is at the root of a great deal of current debate. Joshi and Houtzager’s recent (2012) analysis point out that many proponents of transparency and accountability approaches to improving service delivery assume a relatively confrontational relationship between providers and citizens. In this framing, providers will “misbehave unless constantly watched” (151) by citizens-as-consumers or by a vigilant community seeking to expose corruption and underperformance.<sup>7</sup> Information contributes to this process by emboldening and empowering beneficiaries and other citizens to confront predatory or inept officials. Yet there is another influential view: those who presume “a more trusting, collaborative approach to resolve issues of poor services through collective deliberation and joint problem solving” (152). Proponents associated with this view are associated with efforts at “deepening democracy” (Fung 2001).

These two approaches lead to very different conceptions of how T/A should work.

---

<sup>7</sup> Joshi and Houtzager trace the roots of this perspective to the New Public Management approach, which champions market mechanism to improve public sector accountability and performance (Bately 1999, Joshi 2008).

The confrontational view recommends solutions such as complaint hotlines or citizen charters that help expose corruption and malfeasance. The now iconic success of the *Mazdoor Kisan Shakti Sangathan* (MKSS) in Rajasthan, India is an example of this model. In its early days, that organization helped villagers in Rajasthan to pry budget information out of the hands of secretive governments and organized very public meetings to show the dramatic leakage and theft from the public fisc (Jenkins and Goetz 1999).

The collaborative approach, by contrast, is associated with tools such as the community scorecard that rely on collaborative problem solving between citizens and providers. The Björkman and Svensson (2009) intervention with which we open this paper is a well-known example. In their intervention, community members and health care workers, using information about local clinics and health conditions distributed as part of the intervention, developed action plans to improve the quality of health services and care. Björkman and Svensson (2009) find that this intervention produced large increases in health care utilization, reduced child mortality, and increased child weight.

Whether a generally<sup>8</sup> confrontational or collaborative approach is more likely to be successful is likely to depend on the “willingness” of politicians, policymakers, and front-line providers to make improvements to public services. This willingness, in turn, depends both on the disposition of the individuals involved and the political environment in which they are embedded. The *political environment*—including factors such as formal accountability structures and the degree of political competition—can, depending on the degree of political decentralization, vary between states, districts, or individual

---

<sup>8</sup> The distinction between confrontation and collaboration is not perfect. One consequence of the Uganda intervention, for example, was increased community monitoring of health care providers, a process that could be seen as somewhat confrontational.



villages in the degree to which it incentivizes individual officials and providers to improve public services. And *individuals*—providers, officials, policymakers, and politicians—in the same political environment can vary in their response to those incentives.

The third important consideration is the usefulness of the information itself. Not all transparency interventions provide citizens with information that is useful to them in improving health services, particularly in an environment in which citizens lack the ability to choose among multiple providers of health services. To be useful, transparency interventions must trigger an “action cycle” of changing citizen behavior leading to a productive response by providers.<sup>9</sup> There are four elements of a successful action cycle. The intervention must provide information that is: (1) salient and valuable to users of health services or their allies in civil society, and (2) usable to those citizens, in that it is helpful in changing their behavior or decisions in ways that trigger improvements in governance. Note the implication of these first two elements for the degree to which citizens must *demand* improvements in the service the transparency program is seeking to improve. A transparency program may fail to improve services if users of the service simply do not care about the new information to which transparency gave them access. For the second step in the action cycle to be triggered—for users to take advantage of the new information to select better providers, exert pressure on providers or policymakers, etc.—users must both want the service and be dissatisfied with its current provision. Only then can they be expected to undertake behavior change.<sup>10</sup> For that behavior

---

<sup>9</sup> See Fung, Graham, and Weil (2007) for more on the “action cycle” triggered by effective transparency programs and policies.

<sup>10</sup> This condition may seem obvious, yet perhaps partly for that reason it is often not taken into account in the design of transparency programs: many interventions begin by simply assuming that because a public

change to actually trigger improvements in a service requires two additional steps by the disclosers or objects of the information: they must be (3) sensitive to the information, so that they are affected by its disclosure, and the disclosure of the information must (4) provoke in them a constructive response, so that they act to improve their performance rather than resisting or strategically gaming the disclosure to obscure their underperformance.

Combining these three frameworks:

1. the short and long routes of accountability,
2. the willingness (or unwillingness) of providers, policymakers, and politicians to make improvements, and
3. the criteria for useful information

helps us to understand important contextual factors that influence the path by which transparency might translate into improved services. Together they offer criteria for a successful pathway from the information provided by transparency to service improvement, a set of types of action that service users might take to use that information, and some important considerations in the kind of actions that are likely to lead providers to respond constructively. Here we combine them into five “worlds” of

---

service is failing by some international standard—that, for example, infant mortality is high, or learning outcomes are low—it must be dissatisfying to its users. But this may not be the case. For example, for many citizens in poorer countries, traditional healers are highly preferred to modern clinics for cultural reasons. In many poor countries, basic education has a low rate of return: those who have graduated from primary school do not earn much more than those who spent that time learning traditional skills. Even if this low return is the result of the low quality of schooling—and thus could be expected to increase if users pressed for and succeeded in obtaining improved schooling quality—this hypothetical payoff may not be credible for many citizens. The result is an appearance of citizen “apathy” about the low quality of clinics or schooling that in reality may be quite rational (in cases where the perceived benefit to the service is low) or understandable (in cases where citizens have low expectations for the service—having no experience with anything superior—or a cultural preference for traditional services). In these contexts, a transparency program has little hope of improving a public service unless it can first convince citizens to both want the service and to be dissatisfied with how it is being provided. Only then is it reasonable to expect them to be dissatisfied with its underperformance to the point where they may be willing to do something to improve it.

service delivery: varied contexts in which transparency should lead to service improvement through a different pathway.

The five worlds are pictured in Table 1. In each of these worlds, we hypothesize the role of information and transparency to be different, and therefore that each calls for a different kind of actions and reform strategy. The implication is that the very same transparency intervention applied successfully in one of these contexts may be counterproductive in another.

*Table 1 - Competition, Collaboration, and Confrontation: Five Worlds Facing Transparency Interventions*

		<i>Political Context</i>		
		<i>Competition / Choice</i>	<i>Willing or Accountable Officials (either front-line or policy-makers)</i>	<i>Unwilling or Unaccountable Officials (either front-line or policy-makers)</i>
<i>Accountability Route</i>	<i>Short Route (directly target frontline providers)</i>	1 (market mechanism)	2 (collaboration with frontline providers)	3 (confrontation with frontline providers)
	<i>Long Route (target providers indirectly through policymakers and politicians)</i>	[N/A]	4 (collaboration with policymakers and politicians overseeing providers)	5 (confrontation with policymakers and politicians overseeing providers)

Briefly, the five worlds, and the hypothesized link between information and improvements in health care, are as follows:

1. *Competitive service*, in which beneficiaries can choose among more than one service provider. In this case, transparency can improve outcomes by helping clients make better choices among these providers. Where, for example, a citizen can choose between several public or private clinics, transparency might provide information about the comparative health outcomes of those clinics. This information might be (1) *salient* to patients because they care about their health and find this information credible; (2) *actionable*, in that they can act on that information by choosing a different provider; and (3) *sensitive to providers*, because if providers are funded on a capitation basis, poorly performing clinics will be sensitive to the loss of patients; in turn these providers might (4) *respond constructively* by improving service delivery.
2. *Little competition, but individual public service providers appear willing to engage in reforms*. Here the contribution of information is to enhance the efforts of these providers. Information about actual practices and outcomes may feed short-route collaborative problem solving deliberations between community members and service providers in which both develop more effective delivery practices. Information may also help to document the most critical material needs and make that case via the long route to policymakers and politicians. For a health transparency project, transparency might provide information such as comparative performance data or citizen satisfaction with different aspects of a clinic (perhaps hours, availability of medicines, or treatment experiences). Users of this information might include both clinic clients and health care workers themselves. The information is (1) *salient* to clients because they care about their health and to

frontline workers because they care about the quality of the service they provide. The information might then be (2) *actionable* in triggering a social action component in which the community and providers develop a joint action plans that addresses problems diagnosed by the informational component of the intervention and to which providers are (3) *sensitive*, because it entails a shared sense of commitment and responsibility for improving service. The implementation of this clinic improvement plan would then constitute a (4) *constructive response* from providers.

3. *Little competition and service providers appear unwilling to engage in reforms.*

The third world is more challenging than the second: service providers are unwilling to engage in reform efforts, perhaps because doing so would be costly to their income, privileges, status, or free time. To improve services in this world, the T/A intervention must shift the balance of power between citizens and providers, from a balance in which providers are able to ignore the costs of their recalcitrance for the quality of care available to citizens, to a balance in which providers find it difficult to ignore these costs. In this world, if citizens try to improve service provision directly via the short route, they must do so either by engaging directly with providers to mount pressure for improvements or interventions must be designed to otherwise alter the incentives of these providers. The main contribution of information is to provide clarity about the main areas of underperformance and enhance the capacity of beneficiaries to pressure providers to increase their performance. For example, transparency about health clinic quality might provide information that empowers clinic

beneficiaries relative to providers, such as absentee data, information from audits of materials or finances that reveals leakage, and so on. Beneficiaries might find this information (1) *salient* to them because it reveals that providers are violating social norms and legal requirements; and (2) *actionable*, in that it might enable beneficiaries to act on the information by helping them to organize to protest these violations and advocate for reforms. Health care providers may be (3) *sensitive* to that collective action if it threatens their prerogatives or position, and they may (4) *respond constructively* to those threats by putting forth more effort, properly allocating funds and materials, and so on.

4. *Little competition and service providers appear unwilling to engage in reforms, but policymakers and/or politicians do appear willing to engage.* In this world, even though service providers may not be willing to engage with citizens or respond to citizen pressure, citizens or their champions may engage policymakers at a higher level in the reform project. If such policymakers with jurisdiction over the sector are motivated to improve that sector, information can trigger or reinforce their reform efforts. For example, a transparency intervention that provides new information about how a public clinic or school is underperforming may empower reform-minded policymakers who operate at a level above the direct service providers, such as local political officials or district health administrators. Thus the prime users of information need not be citizens themselves; they may be CSO advocates and policymakers. In this world, the information provided by transparency—perhaps about absenteeism in health clinics—might be (1) *salient and valuable* to civil society advocates who are

seeking to make a case that health care reform is necessary and to develop recommendations for the shape of that reform; (2) *actionable*, in that they can act on that information by incorporating it into their policy advocacy efforts to sympathetic policymakers. Those policymakers, in turn, can use and act on that information by crafting policies such as incentives or sanctions to reduce front-line provider absenteeism. If those policies are well-implemented, front-line providers may be (3) *sensitive* to the new incentives and (4) *respond constructively* by reducing their absenteeism.

5. *Little competition and neither providers nor policymakers appear willing to reform.* In this world—by far the most difficult one for T/A interventions to make a difference—political and administrative elites and service providers are all resistant to reform. The potential contribution of information is thus to mobilize more broad-based social action that will ultimately alter the incentives of policymakers and, thereby, providers. For example, a transparency intervention that provides new information about underperformance of public health clinics and the lack of response by politicians and policymakers to that underperformance might trigger civil society or citizen collective action to incentivize or punish either providers or policymakers and politicians—for example, through social protest or at the ballot box. In this world, such information is (1) *salient and valuable* to citizens and civil society advocates who are seeking to make a case that health care reform is necessary and to develop ways of pressuring reluctant providers, policymakers, and politicians to adopt those reforms; (2) *actionable*, in that they can act on that information by using it

to pressure existing public officials or elect new, more reform-minded officials. Otherwise unwilling officials may be (3) *sensitive* to the public pressure or the possibility of being replaced by more reform-minded officials and thus (4) *respond constructively* by implementing health care reforms and putting renewed top-down pressure on providers to improve their performance.

In Table 2 we sketch these five political contexts of service provision, the differing role of information in each of them, whether the path of action lies on the “short” or “long” routes to accountability, whether this action is primarily collaborative or confrontational, and some exemplary transparency interventions that may produce usable information.



*Table 2 - From Transparency to Accountability in the Five Worlds*

<i>Service Provision Context</i>	<i>Contribution of Information</i>	<i>Accountability Path</i>	<i>Collaboration or Confrontation</i>	<i>Exemplary Transparency Intervention</i>
1 Competition between providers	Inform individual choices	Short Route	N/A	Comparative provider performance rankings
2 Providers willing or have existing incentives to engage in reform	Feed collaborative problem solving	Short Route	Collaboration	Citizen Report Cards or Community Score Cards followed by deliberative community meetings
3 Providers unwilling to reform, community tries to pressure them	Increase pressure and accountability on service providers	Short Route	Confrontation	Community Score Cards followed by naming and shaming, social protest
4 Providers unwilling, policy makers willing to engage in reform	Enable policy-makers to enact top-down reform	Long Route	Collaboration	Social audits, Absenteeism studies, Public Expenditure Tracking Surveys
5 Providers unwilling and policy makers unwilling	Build countervailing power to increase accountability	Short and Long Route	Confrontation	Social audits followed by community mobilization

A few notes on this rubric. First, because it uses factors that may vary within a country, region, and even between communities, a single T/A intervention may find itself dealing with several worlds. For example, the appropriate intervention need not employ the same balance of confrontation and collaboration in all villages and districts; rather, it may combine elements of both, plotting the path of least resistance through the short and long routes of accountability with a mix of collaboration with allies and confrontation with those standing in the way of service improvements. This is the reason that, as we describe in Section III below, we will design our intervention to have a *flexible* social

action component.

Second, the identifying characteristics of each world are complex and thus likely to be far easier to discern for an embedded insider than for even the most experienced outsider. Hence our emphasis—again described in detail below—on *co-designing* our intervention with experienced and locally embedded civil society partners.

Third, note that this rubric is not intended to capture all of the important contextual factors behind the success of Transparency and Accountability interventions. For example, it does not take into account the many contextual elements that may factor into whether citizens are capable of acting collectively to use the information provided by the intervention.<sup>11</sup> Part of the goal of the research design presented in the next section is to allow us to discern additional important contextual factors, which will then be inductively factored into the theory for further testing.

---

<sup>11</sup> That is, for citizens to be able to use information effectively to demand improvements, factors such as citizen agency and the political opportunity structure may also be crucial (Narayan and Petesch 2007, Alsop and Heinsohn 2005). First, citizens may be more successful in using information about service delivery problems—and bringing positive incentives to bear or exerting pressure for improvements—if they can act collectively (Joshi 2008). Given the disincentives to collective action (Olson 1965, Benford and Snow 2000, Gaventa and Barrett 2011, Giugni 1998, McCarthy and Zald 1977, Polletta and Jasper 2001, McAdam 1982, McAdam, Tarrow, and Tilly 2001), where collective action is required a successful T/A intervention may need to facilitate it. Such efforts may be more successful when they are facilitated by civil society organizations that are experienced in mobilizing local citizens, rather than by outsiders (Tandon 2003, McAdam 1982, McAdam, Tarrow, and Tilly 2001). Second, citizens' action—whether individual or collective—may be more successful if citizens can partner at other parts of the accountability chain who are willing to complement citizens' bottom-up activities with top-down or horizontal pressures or incentives (Citizenship DRC 2011, Holland and Thirkell 2009). Again, strong local CSOs may assist this collaboration, as nuanced local knowledge is crucial to navigating the political opportunity structure, finding potential partners and building partnerships with citizen mobilization (Tandon 2003). But local CSOs are not the only prospective partners for accountability: there is a wide literature on citizens working with policymakers (top down accountability), other service providers (sideways or horizontal accountability), and independent media (alternative bottom up accountability) among others can be triggered through citizen actions to engage in accountability work and ultimately increase the likelihood of improving outcomes.

### III. A FLEXIBLE INTERVENTION, CO-DESIGNED WITH LOCAL PARTNERS

In order to take account of—and learn more about—the variety of contexts that transparency and accountability interventions face, we will design, implement, and study an intervention that seeks to improve health outcomes by improving the quality of community health services. The intervention has two important characteristics. First, it provides a standardized informational component combined with a flexible social action component. That is, the intervention will be designed to provide both an increase in information regarding key service delivery problems—the “transparency” component—as well a flexible means through which citizens and communities can use that information to engage providers and policymakers in delivering higher quality health care and overcoming specific health service problems. Second, it will be co-designed with embedded and experienced local partners in order to incorporate their local knowledge about the communities where we will intervene. This intervention will initially be implemented and evaluated in two countries in Phase I of the project, and then in additional countries in Phase II.

The transparency standardized component will focus on uncovering information about issues with service facilities, potentially including the funding or resource level of local clinics, staffing levels, absentee rates, client satisfaction and dissatisfaction, or comparative performance (e.g. the performance of a given clinic compared to the 20 around it).

The second component is social action based upon that information. This social action component is intended to be flexible and adaptive across communities. The range

of social actions may include elements as diverse as community meetings, public hearings, social contracts, co-development of action plans, and petitioning of policymakers. Different social actions, such as action plans, hearings, or policy advocacy, seek to trigger very different underlying political dynamics, such as accountability, collaboration, opposition, or mutual commitment. One of the lessons of the body of research in the T/A field thus far is that the most appropriate mechanism often depends on context. Even within a region, much less a whole country, context varies widely, across districts and individual communities. Thus rather than pre-selecting one strategy—and so risk imposing a particular strategy that is inappropriate in some of the circumstances of our study—our intervention will incorporate a range of such strategies and so allow flexibility, adaptation, and judgment by communities. This flexibility also has a research benefit: it allows us to explore, inductively, the variety of mechanisms that may be triggered in communities of different types or with different characteristics, allowing a deeper analysis of the role of context in determining how transparency translates (or fails to translate) into improvements in health care. We take up this research benefit in more depth in sections III and IV below.

A second distinctive characteristic of our intervention is that it will be co-designed with local partners. This co-design will allow us to build the intervention around the existing literature about which T/A interventions have been effective and which have not, while also allowing us to leverage the experience and local knowledge of CSOs who have a successful track record of working at the intersection of transparency, accountability, and service provision. Those with this tacit knowledge are at an advantageous vantage point for understanding what strategies or set of strategies have the

greatest chance of working in the communities in which we implement the intervention. CSOs, of course, vary in enormously in their capacity and level of focus and embeddedness in local communities; for this project, will work with national-level CSO partners with networks of embedded community-based organizations. The criteria and process for CSO partner selection are in the Appendix.

The remainder of this section explains the structure of the intervention and the co-design process that we will follow to fully develop the details of the intervention.

### *1. The Structure of the Intervention*

In preparation for this project, we undertook a series of CSO surveys and practitioner and academic interviews, as well as an extensive literature review of common T/A interventions that have been evaluated with experimental methods.<sup>12</sup> During this review we noted a few patterns in the information provided in successful interventions:

- Most interventions that focused on service provider inputs (such as absenteeism or financial resources) were successful; most of those that focused on outputs (such as test scores, whose connection to observable inputs may have been unclear to users) were unsuccessful.
- Many successful interventions presented comparative information that allowed users to see how their providers were performing relative to other villages or to

---

<sup>12</sup> Reinikka and Svensson (2005), Pandey et al. (2007), Olken (2007), Hastings and Weinstein (2008), Banerjee, Duflo, and Glennerster (2008), Björkman and Svensson (2009), Andrabi, Das, and Khwaja (2009), Pandey, Goyal, and Sundararaman (2009), Banerjee et al. (2010), Muralidharan and Sundararaman (2010), Keefer and Khemani (2011), Pradhan et al. (2011a), Glewwe and Maïga (2011), Reinikka and Svensson (2011), Duflo, Hanna, and Ryan (2012), and Lieberman, Posner, and Tsai (2012).

national standards; nearly all studies of unsuccessful interventions presented just absolute information (for example, test scores).

- Many successful interventions presented both *objective* and *subjective* information (for example, both medical stocks and citizens' perceptions of waiting times at clinics).

In addition, our review showed that many T/A interventions do not prescribe a specific type of action that should be employed after information is collected. In theory, this is a good practice; our own research and theorizing imply that different contexts and types of information are better suited to different types of actions, such that a “one size fits all” approach is unlikely to be consistently effective. However, in reality, the fact that many interventions provide no guidance on community or stakeholder action means that the intervention often ends at the point of information collection and distribution.

Several additional patterns emerge from our review:

- Many successful interventions include a *collective action* component. Interventions that do not try to relax a collective action constraint seem less likely to be successful.
- Those interventions that *define specific actions* for citizens or stakeholders to take are more often effective than those that merely disseminate information.

We sought to evaluate an intervention that builds from these lessons, as well as meeting several additional criteria:

- *Potential for adaptability to different problems and contexts.* We note above that one potential reason for conflicting evidence regarding the effectiveness of some types of T/A interventions is that existing interventions may be effective in some

contexts but not others. While most T/A interventions have a theory of change—in which information about the service delivery problem (transparency) triggers individual or social actions to hold responsible parties to account (accountability) leading ultimately to service improvements—those theories of change by and large build in strict assumptions regarding how certain agents act and react. These assumptions are far more likely to hold true in some contexts than in others. Thus we sought an intervention with a theory of change flexible enough to adapt to different contexts.

- *“User friendly” for CSOs.* Because our goal is an intervention that has the potential to be adopted and adapted easily by CSOs, we sought to identify an intervention with relatively low technical or resource requirements.
- *Likely to trigger health improvements at the community level, through either top-down or bottom-up approaches.* Rather than pre-define “bottom up” or “top down” mechanisms, we sought an intervention that is likely to impact individual level outcomes at the community level but could allow for agents of change either at the community level or at a higher level, such as regional public officials or politicians.

While we identified intervention types that met some of these criteria, no single intervention met all these criteria described above. Instead we are developing a hybrid approach, the “Community Score Card – Plus.”

The information component of the T/A intervention is based on the common Community Scorecard (CSC) approach. The CSC has proven effective in the health sector in Uganda (Björkman and Svensson 2009) and has the advantage of targeting both

community members and front-line service providers. Its two initial steps use information that is both subjective and objective, focuses on inputs (as well as outputs), and is comparative. Following the compilation of facility scorecards dealing with both inputs and outputs,<sup>13</sup> both providers and community members complete subjective assessments on service performance.<sup>14</sup>

The final step in a traditional CSC is the development of an action plan in which community members and providers commit to taking certain steps to improve the service. This component of the CSC does not allow sufficient flexibility in approach to follow the three lessons outlined above. Thus our CSC-Plus technique will replace the final step with the creation of a flexible social action plan that can respond to context. While all components of the CSC-Plus technique will be co-designed with our local partners, the co-design process will put particular emphasis on the components of this flexible social action plan, based on their local knowledge of effective means of translating information (transparency) into improvements in service delivery (accountability). The resulting plan will form a social action “menu,” from which individual treatment communities,

---

<sup>13</sup> The purpose of the scorecard is to allow civil society to collect information on both inputs and outputs in facility-level service delivery (in this case, health care) to identify areas in which services are not being provided effectively. Specific components of the scorecard can vary, but they generally contain information such as stock levels of medicines, facility conditions, the presence of providers, utilization rates (how often community members seek preventative and curative services) and in some cases, indicators of health status of the community in the catchment area of the facility. A common scorecard is developed for a large sample of facilities, allowing collected information to be compared across different sites. For the purposes of our intervention, the research team and implementing partners would develop a facility scorecard with a subset of questions that would be comparable across not only facilities but also countries, allowing us to test the generalizability of this intervention.

<sup>14</sup> Information from subjective assessments is often invaluable in T/A interventions to answer questions of why breakdowns identified in the scorecards may be occurring or how the problems might be solved from the perspective of providers or beneficiaries. In addition, these assessments can uncover discrepancies in perceptions that might lead a particular collective action message or technique to fail. For example, a scorecard may identify widespread problems with health provider effort; however if the user assessment shows that citizens are largely satisfied with the level of effort from doctors, this indicates that an advocacy campaign to encourage citizens to request government-led monitoring of health provider effort is likely to fail, as citizens may not see a need for improving service provider effort.



following the completion and distribution of the facility scorecard and subjective assessments, will be able to select one or more actions that they find to fit best with their community, its context, and the particular service delivery problems it faces. For example:

- A community operating in a context with limited choice of providers but providers that are open to reforms (our world 2) might select a collaborative problem-solving meeting with service providers to brainstorm ways to improve service delivery and trigger horizontal (peer-to-peer) accountability;
- A community operating in a context in which service providers are perceived as unwilling but district government officials are reform-minded (our world 4) may select a social action in which they request a citizen charter signed by district officials to trigger top-down accountability; or
- A community operating in a context in which neither providers nor local officials are willing to make changes to improve service quality (our world 5) may select a social action in which they work with local radio to “name and shame” providers and government officials who are not performing well.

Each of these social actions would be associated with a different theory of change, in which the particular action leads to improved service delivery and ultimately improved health outcomes.

## 2. *Co-Design*

The co-design process of the “CSC Plus” intervention will be engineered to allow as much consistency between the interventions pursued in the two Phase I sites as possible, without sacrificing the ability of each site to utilize mechanisms and triggers that are most likely to be successful in their particular contexts. The first step will be to identify with our CSO partners the specific problems the intervention will target and the health outcomes that it will seek to improve,<sup>15</sup> as well as the area of the country where they will implement the intervention. We will then work with the implementing partners to develop

1. a common set of informational components—the pre-intervention diagnostic, the facility report card, and the service assessments by users and providers—that can cover a set of potential problems in health service delivery that is broad enough to take account of the specific service delivery problems that the CSOs deem to be important across both sites.
2. a diagnostic tool—the pre-intervention diagnostic—to help identify which of the “five worlds” of service delivery are represented within our intervention areas. This diagnostic tool will help the CSO identify the combinations of information collection, dissemination, and social action strategies likely to be successful in a given “world,” as well as the relevant worlds represented within the intervention area—so that we can ensure that we have sufficient variation in contexts for the

---

<sup>15</sup> For example, the intervention might focus on a specific problem such as absenteeism at health facilities, or on a broader measure of the quality of services at a health facility including, for example, the presence of health workers, availability of resources, and the quality of care and treatment offered. Both the objective of the intervention and the health outcomes measured will be consistent across all of the communities in the evaluation.

purposes of our evaluation. For example, in Tanzania, we might decide to work in two neighboring districts one of which is dominated by the political party in power and another in which the opposition party is more competitive, making district officials and front-line service providers more sensitive to citizen demands for service improvement.

The social action menu resulting from this co-design process will be designed to include actions from both Phase 1 CSOs, so that it, like the informational components of the intervention, is consistent across the two Phase 1 sites.

The interventions and associated data collection will be piloted at a small scale to identify any potential bottlenecks and issues, and allow refinement of the intervention and evaluation approach, prior to full implementation.

#### IV. A MULTI-STAGE, MIXED-METHOD RESEARCH PLAN

Our project is designed around 9 research questions:

1. What is the effect of our flexible, co-designed T/A intervention on the quality of health service delivery?
2. What is the effect of the intervention on individual health outcomes?
3. If there are significant effects, what are the mechanisms through which these occur?
4. What is the role of context in shaping or determining these mechanisms?

5. What are the implications of the intervention for citizens' perceptions of empowerment and efficacy, both within communities and between communities and the state?
6. How can the design of the intervention itself be further improved?
7. How does the intervention stack up against alternative health interventions (e.g., drug, vaccine, or diagnostics) or other T/A interventions in terms of relative cost-effectiveness?
8. Does our intervention have similar effects beyond the initial countries where we implement it?
9. Can our theoretical insights about mechanisms and context be generalized?

To explore these questions, we have designed a two-phase, mixed method research program. In developing the program we were guided by twin goals: to assess the causal effects of our intervention as rigorously as possible, as well as to gain a deep and practically-applicable understanding of the roots of these effects and their likely generalizability. These goals led us to an approach that closely integrates randomized controlled trials with a variety of intensive qualitative methods.<sup>16</sup> Our program combines careful selection of initial cases; an initial phase in each of these cases in which we evaluate our intervention with both a randomized controlled trial and an integrated and extensive qualitative evaluation to understand mechanisms and the role of context; a cost-effectiveness analysis; and a second phase, which explores the generalizability of both our first phase results and the theoretical implications of them for mechanisms and the role of context.

The following two sub-sections outline the two phases of the research program.

---

<sup>16</sup> For example, Stern et al. (2012), King, Keohane, and Verba (1994).

## 2. *Phase I*

Following the co-design process, the first phase of our research project involves evaluating intervention effects, both on health care quality and outcomes, and on community power relations and dynamics. Phase I has two goals. First, it is our primary investigation into the efficacy of transparency and accountability interventions for improving health care quality and outcomes. Second, it is the first step in building an empirically verified theory of the impact of transparency and accountability interventions.

Because Phase I is partly a theory-building phase, we have selected two Phase I sites, one in Africa and one in Asia, that provide us with variation on potentially important dimensions of context: geography, history, and culture; and development and democratization. In contrast to the general method of selecting sites for theory-testing, which often seeks to control for contextual factors (by, for example, randomization or selection of sites that have many contextual factors in common), sites for theory-building benefit from maximal breadth of contextual factors. We sought sites that allow as many potential contextual factors to play a role as possible, so that we can observe their implications and use them in developing our theory—particularly in developing hypotheses for which contextual factors play the largest role in the success of our intervention. (These hypotheses will be further evaluated in Phase II.)

On a more practical level, we also selected sites with strong civil society partners at the country level: CSOs with the deep local knowledge necessary to navigate the particular contexts of citizen action and the political opportunity structure that, as we note

above, are likely to be key to the impact of our intervention. In addition, we sought local partners who have the motivation to undertake in this work for their own reasons: partners who are interested in engaging with us in the long-term process of developing and evaluating a T/A intervention, who are invested in the evaluation itself—perhaps because it will provide insights that will help them improve their own work. Finally, we sought partners who are embedded in networks that will allow them to spread the research findings among a wider range of practitioners. This process led us to the selection of Tanzania and Indonesia as our Phase I sites. The case selection criteria and process are detailed in the Appendix.

In each of these two sites, we will conduct a Randomized Control Trial, with between 50 and 100 control communities and 50 and 150 treatment communities, to assess the extent of the impact of T/A interventions on health service quality and individual health outcomes. The RCTs are intended to provide answers to the first two of our research questions:

1. What is the effect of the Phase I interventions on the quality of health service delivery?
2. What is the effect of the Phase I interventions on individual health outcomes?

The precise outcomes we will measure can only be determined once we have co-designed the intervention with our local CSO partners (see Section III). But we anticipate measuring outcomes in three broad categories: indicators of citizen monitoring and empowerment; measures of facility performance (treatment practices and management); and individual level health output and outcome measures. Table 3 provides illustrative examples.

*Table 3 – Categories of Potential Outcome Measures*

<i>Indicators of Citizen Monitoring and Empowerment</i>	<i>Facility Performance (treatment practices and management)</i>	<i>Individual level health output and outcome measures</i>
<ul style="list-style-type: none"> <li>Observational indicators as to whether communities put in place monitoring techniques such as suggestion boxes, numbered waiting cards to make wait times and process more predictable for patients, or other measures to monitor staff absenteeism</li> <li>Measures of citizen perceptions of empowerment and efficacy</li> </ul>	<ul style="list-style-type: none"> <li>Patient waiting time</li> <li>Staff absence rate</li> <li>Indicators of the physical condition of facilities</li> <li>Drug stock levels/stock-outs</li> </ul>	<ul style="list-style-type: none"> <li>Immunization rates</li> <li>Birth rates</li> <li>Pregnancy rates</li> <li>Weight-for-age measures</li> <li>Number of births attended by a skilled attendant</li> </ul>

Because we will intervene at the facility level, the unit of randomization will be the catchment area of the facility, defined as the community living within a 5 km radius of the facility. This area is large enough to avoid spillovers and small enough to allow for reasonably sized samples. (The precise area will be determined after we have a more precise sense of the intervention design in each site and of the relevant geographic units in each of the countries.) We will collect data at each health facility and from a random sample of households living in the defined catchment area of the facility. For each of our Phase I evaluations, we will conduct baseline and endline data collection, as well as midline data in at least one of the sites (these midline data will be used in the selection of sites for Phase II; see subsection 2 below). Program impacts of Phase I interventions will be estimated both on average and separately for key sub-groups, such as women, the extreme poor, and members of other disadvantaged or disempowered groups.

Following the intervention co-design process, we will produce an impact evaluation design memo that will detail the key design parameters of the RCT design, including: outcome measures, unit of random assignment, sampling and data collection

plans, and sub-groups to be examined. This will provide the blueprint for the RCT implementation as well as minimizing the possibility (or even the appearance of the possibility) of identifying spurious statistical relationships *ex-post*.

Our intervention seeks to understand not only *whether* T/A interventions can improve health care quality and health outcomes, but also *how* and *in what conditions*—all in an effort to develop a robust theory of T/A’s impact that can be a guide to future practitioners and scholars. Thus we will supplement these quantitative methods with a series of qualitative components to allow investigation into research questions 3 – 6:

3. If there are significant effects, what are the mechanisms through which these effects occur?
4. What is the role of context in shaping or determining these mechanisms?
5. What are the implications of the Phase I T/A interventions for citizens’ perceptions of empowerment and efficacy, both within communities and between communities and the state?
6. How can the design of the intervention itself be further improved?

Answering these questions will involve gathering two additional types of data: intensive qualitative data from a smaller sample of communities, as well as supplementary data from all treatment and control communities.

First, we will collect intensive qualitative data on a stratified subset of 20 communities—10 treatment and 10 control communities. The goal will be to gain a deeper understanding of the existing context, the implementation of the intervention, and its implications for health outcomes and the broader political and social structure. In these case studies, our rationale and approach in the case studies for each of the research



questions above is as follows:

- *Mechanisms (question 3)*: We will process-trace the mechanisms underlying any effects, in order to explain the reasons for the results and to account for differences or similarities in these mechanisms, and their effects, across contexts.
- *Context (question 4)*: We expect to see different strategies pursued through the social action component of our intervention in different communities. Should this divergence of mechanisms prove to be the case, our qualitative work will seek to identify sub-groups of communities (delineated, for example, by factors such as prior level of social organization or social capital, education, ethnic ties, community fragmentation and conflict, configurations of local and traditional power, and providers or policymakers willing to engage in collaborative problem solving) within the treatment communities that use the information provided by the intervention in similar ways. By identifying the main characteristics that delineate when some kinds of social action are more appropriate than others, we will seek to further develop the preliminary theory outlined in Section II of the role of T/A interventions in different contexts. (We will further verify this theory in the second phase of the project.)
- *Community dynamics and power relations (question 5)*: We will examine changes in political activity and perceptions of citizen efficacy and provider or policymaker willingness. This is a critical set of variables to consider, not only because it helps us to better understand the mechanisms through which the intervention works but also because it will allow us to examine the implications of the intervention on citizen empowerment. By combining our analyses of these

changes with process-tracing of the mechanisms triggered by the intervention, we will develop an understanding of the interaction of the intervention with existing power relationships among the treatment communities and between the communities and the state. We will also compare these dynamics with those in the control communities, to provide insights into whether the intervention is likely to have empowered or disempowered community members—particularly traditionally disenfranchised groups such as women and ethnic minorities—or was irrelevant to the structure of power.<sup>17</sup> By examining similarities and differences in the responses of state institutions and processes across our case-study communities, we can gain insight into how our intervention interacts with 1) idiosyncratic state relationships and dynamics with specific communities; 2) state-society relationships that are common to communities with certain characteristics; and 3) state-society relationships that are general, at least in the region we are studying. The third category of insights can allow us to understand how a T/A intervention like ours interacts with the state in two very different countries like our Phase I countries—an important concern for donors interested in whether governance efforts like T/A supplement, undermine, or are irrelevant to the state’s own efforts.

- *Intervention design (question 6)*: Finally, we will aim to further improve the design of our intervention by examining the success or failure of the implementation of our intervention across our Phase I treatment communities. By looking at how information is produced and used in the initial interventions, we will refine the quality of the intervention in subsequent interventions. For

---

<sup>17</sup> See, e.g., Alsop and Heinsohn (2005), Holland and Thirkell (2009)

example, we will look at how the information produced by the intervention is received by beneficiaries and communities, with an eye to identifying other information that would be more meaningful, salient, and actionable.

Data for these 20 case study communities will be collected primarily by two on-the-ground observers, trained in ethnographic methods, who will live nearby for the entire multi-year period of the intervention. These observers will keep a close watch on the implementation of the intervention and the dynamics that it triggers. They will collect data primarily with four research techniques: focus groups, process-tracing, trusted local informants, and direct observation.<sup>18</sup> Data collection will focus on the nature and organization of the existing health care system at baseline and over the course of the intervention, beliefs and practices of providers—including the perceptions and biases behind their practices—and power relationships between and among providers, officials, and community members. Case studies will rely both on subjective assessments of providers, officials, and community members, as well as objective data that can be gathered from clinics (who was treated, how, reactions to complaints, etc.) and community meetings (who sat where, who spoke). Case studies will also examine “emergent” sub-groups and mechanisms, which can then be incorporated into endline surveys.

Analysis of data from these communities will proceed as follows. For each community, all data—from surveys, focus groups, existing community level data, and interviews—as well as the material from direct observation and existing secondary sources will be combined into a detailed case study of the community. The 20 resulting

---

<sup>18</sup> Our approach has parallels to ethnographic work that is increasingly common in research on health services in the US and Europe (Savage 2000).

case studies will describe the context prior to the intervention (e.g. social organization, education, ethnic ties, community fragmentation or conflict), and then process-trace the mechanism triggered by the intervention; the response of service providers as well as the response of local elites and state institutions; community members' experiences of the mechanism and institutional response; and perceptions of changing citizen efficacy and power relations. Responses of local informants and community members (both individually and in focus groups) about their perceptions of why different elements in the mechanism occurred and why the whole process proceeded as it did will be compared to direct observation, in an effort to triangulate the elements in the account and assess the relationships and power dynamics underlying the intervention and the particular social actions chosen by different communities—including whose perceptions of the main problem with service delivery are recognized, who takes responsibility for the response and whom that response seeks to serve, and whether the response seems to differentially affect certain groups, such as women or traditionally disadvantaged groups. In the control-cases, the narrative will seek to identify any changes in citizens' perceptions of efficacy and power relations, as well as meaningful changes by service providers and by state institutions, for later comparison to the experience of the treatment communities. These individual case studies will be made publicly available so that practitioners can learn from the experience of actual communities implementing these interventions and scholars can check our interpretations.

To examine broader questions of mechanisms, context, and implications for empowerment, the 20 case studies will be analyzed as a whole. First, in the treatment communities, this analysis will examine the sort of mechanisms that were produced by

the intervention, and whether these converged or diverged across communities with different contextual characteristics: our “five worlds” of service provision in addition to the wide range of other characteristics that might condition how T/A will work (e.g. social organization, education, ethnic ties, community fragmentation or conflict). The mechanisms in each case study will be classified as one of the hypothesized mechanisms the team developed prior to the intervention (e.g. collaborative or oppositional strategies; see Section II) or as a new mechanism we did not anticipate. Should we see different mechanisms pursued in different communities, the analysis will seek to determine which community-level factors were common to communities that pursued each mechanism, and subsequent investigations—through additional focus groups and interviews with community members and key informants—will triangulate around plausible causal pathways from particular community contextual factors to a particular mechanism. The question is: what are the most important contextual factors that led community *X* (given its set of characteristics) to pursue mechanism *A* and community *Y* (given its set of characteristics) to pursue mechanism *B*? In other words, the analysis will try to establish plausible context-mechanism pairings. For example, it may be the case that clusters of high-education or otherwise advantaged communities are more comfortable pursuing collaborative strategies, while those of low-education favor protest or other oppositional strategies. Alternatively, the effect of transparency may be mediated through local power structures or cleavages of community conflict. These context-mechanism pairings will then feed back into the development of our broad theory of T/A’s impact for further verification in Phase II.

In addition, the analysis of the case studies will focus on changes in community empowerment and interactions with local elites and state institutions. Any changes in perceptions of empowerment by community members—particularly traditionally disenfranchised groups such as women and ethnic minorities—will be compared across treatment and control communities so as to separate the changes that can be process-traced back to the mechanisms triggered by the intervention. Separately, the analysis of the treatment communities will examine patterns of response to the intervention by elites or state agencies—responses that might be both reinforcing of the intervention (for example, local elites talking up its successes at bringing the community together or improving services; state agencies adopting or spreading more widely some of the solutions developed as part of the social action component of the intervention) or undermining (for example, state agencies finding ways to undermine new accountability instruments set up as a result of the intervention, or allowing increasing leakage of some resource whose procurement process was previously relatively uncorrupt to make up for increasing monitoring of procurement of another resource as a result of the intervention). Responses in the case study communities will be classified and then grouped into three categories: 1) responses that were entirely unique to a particular community; 2) responses that were common to all communities with either certain characteristics or that pursued certain mechanisms; and 3) responses that were common across all treatment communities. These will provide insights into interaction of the intervention with the prevailing institutions and patterns of interaction in three ways: 1) responses that can be predicted across all communities in the country (across, for example, all of Tanzania or

Indonesia); 2) responses that may be expected when a community in the country has certain characteristics; and 3) responses that are unpredictable.

Finally, the intervention-implementation experiences across the case studies will be examined for patterns that we can use to improve the intervention, both for our own Phase II analyses, and for wider dissemination to any practitioners looking to implement similar interventions. We will analyze the cases for patterns in how information is produced and used in the initial interventions, how the information produced by the intervention is received by beneficiaries and communities, and the particular forms of social action it triggers (e.g., deliberation, collective action, policy-reform). This analysis will help develop a granular understanding of what sorts of information are more salient and actionable for individuals in various situations as well as how to refine the social action prescriptions.

The intensive qualitative data from the 20 case study communities will be supplemented by additional data gathered in the surveys fielded in the RCT impact assessment. These surveys will be designed to provide the same sort of information as in the 20 case studies, on important community characteristics; mechanisms triggered by the intervention; perceptions of community dynamics and power relations; and the suitability of the intervention design. For example, surveys of all treatment and control communities will assess prior level of social organization or social capital, education, ethnic ties, community fragmentation and conflict, configurations of local and traditional power, and providers or policymakers willing to engage in collaborative problem solving. Surveys will also seek to assess the mechanism triggered by the intervention; the response of service providers as well as the response of local elites and state institutions;

community members' experiences of the mechanism and institutional response; and perceptions of changing citizen efficacy and power relations. We will supplement these with data from existing sources (for example, existing surveys and census and other official data) on community and facility characteristics, such as socioeconomic and demographic characteristics and the competitiveness of elections for political offices. These data, while necessarily more cursory than the intensive case studies, will help to verify the findings from the 20 case studies in the broader sample of communities.

Table 4 summarizes the types and purposes of the data gathered as part of the investigation into research questions 3 – 6.



*Table 4 – Data Sources, Samples, and Purposes for Research Questions 3 – 6*

<i>Data Sources (illustrative)</i>	<i>Sample</i>	<i>Purpose of Data Collection/Examples of Data Items</i>
<i>(1) All communities</i>		
Baseline Survey	Households	Obtain information on key variables to be used for sub-group impact analyses, increase the statistical power of the design, confirm the equivalence between treatment and control groups, learn more about the target population of each intervention, including their assessments of their efficacy and community relations, and to generate report cards that will be distributed to citizens in treatment communities.
Facility Records	Health Centers	Collect facility-level data (e.g. patient utilization, health worker presence, hours of operation, etc.)
Community-level data	Communities	Gather community-level data from existing sources to understand socioeconomic and demographic conditions as well as political competitiveness (where relevant).
Endline Survey	Households	Measure key outcomes (e.g. health worker attendance, disease prevalence, etc) and data on potential mechanisms behind the effects
<i>(2) Stratified sub-sample of communities</i>		
Focus Groups	Stratified sub-sample of 10 treatment and 10 control communities	Gain a qualitative understanding of the context in which interventions are being implemented (e.g. key implementation issues, perceived benefits and costs of the intervention, willingness and effectiveness of providers, officials, and politicians, political activity, and perceptions of citizen efficacy and community power relations)
Process-tracing	Stratified sub-sample of 10 treatment communities	Assess qualitatively the effects of the intervention on important community actors, the mechanisms the community decides to pursue in its action plan, the factors that led the community to select these mechanisms, and the system-level responses to the intervention.
Trusted local informants	Identical sample to in-depth case studies	Gain informed assessments of about the political, social, and service structure and about the existing dynamics of conflict and cooperation around services to supplement the process-tracing in the in-depth case study communities
Direct field observation	Identical sample to in-depth case studies	Gain additional insights through ethnographic techniques to monitor the implementation of transparency interventions in the in-depth case-study communities and the consequences that result from them.

The final research question in Phase I is whether T/A interventions can add value to traditional health sector interventions:

7. How do T/A interventions stack up against alternative health interventions (e.g., drug, vaccine, or diagnostics) or other T/A interventions in terms of relative cost-effectiveness?

We will address this question by conducting a cost effectiveness analysis in each

of the Phase I sites. Estimates from this analysis will be compared to those of other TA interventions and to typical technical interventions aimed at improving health outcomes.<sup>19</sup>

## *2. Phase II*

The second phase of our work, which will unfold over the last two years of the project, is designed to explore our last two research questions:

8. Does our intervention have similar effects beyond the two regions investigated in Phase I?
9. Do our theoretical insights about mechanisms and context apply outside the two regions investigated in Phase I?

The Phase II research design will depend on the research findings from Phase I: in particular, whether Phase I reveals positive impacts and whether the intervention triggers the same mechanism or different mechanisms across different contexts. Depending on these results from Phase I, Phase II will either involve a third RCT or a set of four smaller studies in additional sites.<sup>20</sup> Figure 1 shows the key design choices of Phase II depending

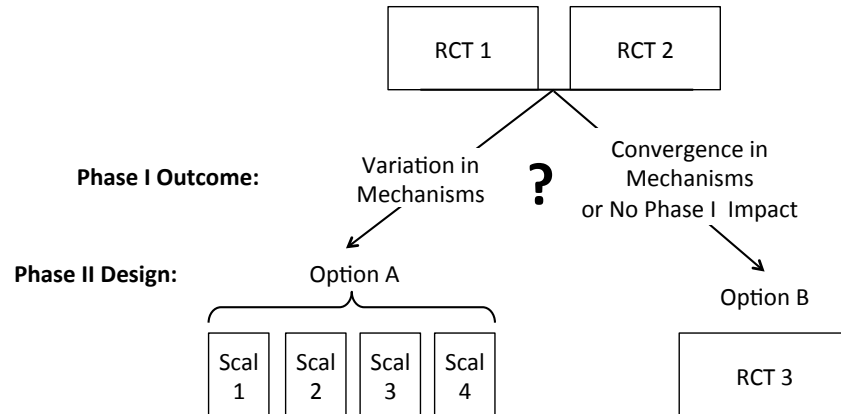
---

<sup>19</sup> The difficulty with this kind of analysis is that T/A interventions seek to do more than merely improve individual-level development outcomes; they also seek to improve citizen agency, empowerment, and other political variables. We will therefore be careful to use our analysis to see whether T/A interventions compare favorably (e.g., are they in the same cost effectiveness range?) as alternative technical (e.g., medical) interventions, while also noting that this is only one possible benefit of a T/A technique. If our studies show impact on political variables, we would be able to indicate that these were additional benefits. One way to deal with interventions that have multiple outcomes is to conduct a cost-benefit analysis rather than a cost effectiveness analysis. Though T/A interventions have multiple outcomes, we have decided not to do a cost-benefit analysis for several important reasons, including difficulty of monetizing certain impacts of these interventions (such as empowerment and citizen agency) and the fact that cost effectiveness estimates are more prevalent for a variety of interventions that we would like to use as comparators.

<sup>20</sup> In order to allow time to launch and complete Phase II by the end of the project in year 5, we will use midline data and analysis from Phase I studies to inform the structure of Phase II.

on the Phase I results; these options are detailed below.

*Figure 1 – Phase II Design Options*



The first possibility is that a common pattern of information use (e.g. comparative facility performance) and social action (e.g. community action such as protests or petitions) emerges from the Phase I as a prime driver of service delivery improvements, and that this pattern holds across contexts—across communities with very different characteristics in both our two Phase I sites. If such convergence results from Phase I, then the challenge of Phase II is to validate that intervention across a broader range of social, economic, and political contexts. In order to meet this challenge, we will develop a unified Phase II intervention (depicted as Option B above) that is a simplified form and so less demanding and less flexible than the intervention fielded in Phase I. We will then conduct a third RCT of that intervention in a sample of communities covering a more diverse geographic area within one country. This Phase II design will test scale—through geographic reach—and generalizability—through replication and implementation in an entirely new range of socio-political contexts than Phase I.

A second empirical outcome of our Phase I trials is that we find no significant

effect on health outcomes. In the event of this outcome, our research strategy will still allow us to observe the mechanisms through which the interventions should work. Thus, even if the intervention does not have the ultimate positive effect on health outcome measures, we will review the process data to develop a new intervention based on information about what mechanisms were triggered and where the intervention failed. Ultimately, if the Phase I trials produce null effects, we will conduct a third RCT, testing the new intervention developed based on the lessons from the earlier phase.

A third possibility is that our Phase I trials reveal significant impact (service delivery improves on average) but different mechanisms or pathways to that impact in different contexts. That is, our qualitative research may show that certain contextual features—such as those in our “five worlds” rubric of service delivery, or other factors such as the prior level of education or social capital—are particularly important for the success of a given mechanism. If such divergence, rather than convergence, is the result of Phase I, then the challenge of Phase II is to validate the Phase I understanding of why some kinds of T/A intervention are suited to some contexts but not others (depicted as Option A in Figure 1). Option A involves a series of four smaller studies, in four new sites, that utilize mixed methods (qualitative case study and quantitative analysis) to verify: 1) whether our we have correctly identified the important contextual factors, and 2) whether, under these conditions, our T/A intervention is associated with improvements in service quality.

In contrast to Phase I, where the selection of sites is driven by the requirements of theory-building—and thus our case-selection objective is to choose sites that provide variation across a wide range of potentially important characteristics—our objective in

Phase II is initial theory verification. Thus the appropriate case-selection method for the four Phase II sites is different: we will seek cases that vary on key dimensions that our theory and Phase I results point to as important and that vary on as few other characteristics as possible. For example, suppose we find that our rubric was right that mechanisms differ by political willingness—in particular, say, the degree of political competition. Specifically, suppose that we find that communities in which there is a high degree of political competition converge on collaborative strategies while those with a low degree of political competition converge on more oppositional strategies. In this case, one of the criteria for selecting Phase II sites would be two additional clusters of communities in two separate sites that vary on their level of political competition and on as little else as possible (e.g. culture, level of economic development, history and geography, etc.). If we implement the intervention in both and find that, as hypothesized, it triggers collaborative problem solving in the site with high levels of political competition and oppositional strategies in the site with low levels of political competition, this gives us greater confidence that our Phase I results and subsequent theorizing has identified a more general tendency of a particular contextual characteristic (political competition) to influence *how* our T/A intervention has an impact (by affecting whether communities are likely to pursue collaborative or oppositional strategies).

The four Phase II case studies will be selected to explore the role of two such contextual factors. (We can explore variation in additional characteristics that are likely to vary at the community level—such as education—with our sampling of communities in the Phase II sites, as described below.) For example, we might find in Phase I that, in addition to level of political competition, communities varied in the mechanisms they

chose according to the degree ethnic homogeneity. In this case we would select four cases that varied on “high” and “low” levels of political competition and ethnic homogeneity and on as few other characteristics as possible. Figure 2 presents the case-selection methodology for this hypothetical scenario.

*Figure 2 - Case selection for Phase II for two hypothetical factors*

		Context 1: Political Competition	
		High	Low
Context 2: Ethnic Homogeneity	Low	Site 1	Site 2
	High	Site 3	Site 4

With no budget constraint, we might seek to measure impacts in these contexts with additional RCTs, which would provide the most definitive evidence of generalizability of our Phase I findings. However, given the high cost of RCTs, such a large package of RCTs would not provide good value for money. Instead, we would implement a series of smaller scale studies, which can explore basic mechanisms and lay the groundwork for a more focused package of RCTs. These smaller Phase II studies will focus on 1) whether the observable implications of the expected mechanisms were in fact observed in treatment communities and 2) whether the interventions were associated with perceptions of improvements in service quality—such as community perceptions of increasing usage, satisfaction, and provider effort or more readily available medical resources—that can be triangulated with the result of other methods such as key

informant interviews and direct observation (e.g. of provider effort and absenteeism)<sup>21</sup>:

1. *Assessment of mechanisms*: We will assess whether the same mechanisms that were triggered in Phase I were also triggered in Phase II for the same kinds of communities. In addition to the variation possible in our case-selection (for example, to explore broad factors such as ethnic homogeneity and the degree of political competition), we can also explore community-level variation on additional characteristics that our Phase I studies reveal to be important, such as level of education: if Phase I reveals that communities with relatively highly-educated citizens were successful in implementing collaborative strategies whereas communities with relatively low education employed more oppositional tactics such as protest, then our sampling in Phase II will allow us to examine some communities with highly educated citizens and others with less highly educated citizens. We would then assess whether the mechanisms elicited in Phase I for each of these two groups (collaborative or oppositional) also take place in each of the four new contexts of broader factors (such as the level of political competition or ethnic homogeneity) which should also play a role in the kind of mechanisms our T/A intervention will trigger.
2. *Assessment of quality of health services*: We will assess the extent to which the quality of health services seems to have improved over the course of the intervention. We will not interpret differences in quality of services as the impact of the interventions (as there are many potential confounders) but this assessment

---

<sup>21</sup> These studies would be not intended to be substitutes for RCTs in estimating the impacts of an intervention, a point that will be stressed in all products from the Phase II analyses.

will allow us to assess whether citizens perceive any improvements in the health services that they receive.

Briefly, the methods employed in each of these smaller scale studies would be as follows. In each site, we will choose a stratified random sample of 30 communities that vary in contextual dimensions that were assessed to be key during the qualitative work in Phase I. The studies will involve collection of qualitative and quantitative data.

The qualitative data will be collected through focus groups and through interviews with providers, local officials, community members, and key informants, the key objective of which will be to assess whether the mechanisms triggered by the T/A intervention were consistent with the ones triggered in Phase I.

The assessment of the quality of the health services will rely primarily on data from two sources: 1) a household survey of a random sample of 30 households, in which households will be asked questions about issues such as their usage of health services, choice of facility, satisfaction with health services, perception of change in the quality of health services over time; and 2) direct observation of health care providers to directly gather data on issues such as absenteeism rates and availability of supplies.

Analysis of mechanisms and context will follow the same approach as in Phase I, in which evidence of mechanisms comes from triangulation of information from the various data-gathering techniques, and process-tracing is used to link the mechanism to the intervention on the one side and to any perceptions of change in service quality on the other. Likewise, the impacts of the intervention will be estimated with the same approach as in Phase I. If Phase II involves four smaller studies the quantitative results will be purely descriptive (i.e. will not seek to make causal claims) and will reflect the sampling



uncertainty through the use of standard statistical tools (standard errors and/or confidence intervals). We will seek to increase the validity of the findings by closely integrating the quantitative and qualitative work. For example, if key informants in the qualitative work suggest that the quality of health service has improved since the community started to monitor the health facility, the household survey can be used to assess if a representative sample of households perceive an improvement in the quality of health services and in which type of health services. Then the facility survey can be used to assess the key components of the health services seem to have improved over time.

A central objective of the Phase II analysis is to provide initial confirmation or disconfirmation of the theoretical insights from Phase I, with the overall goal of producing an empirically verified “theory of T/A’s impact.” There are four potential outcomes of Phase II, each with different implications for the development of the theory and for further work on the impact of T/A once our project has ended:

1. The same contextual factors triggered the *same mechanisms* as in Phase I, and evidence indicates *an improvement* in the quality of health care;
2. The same contextual factors triggered the *same mechanisms* as in Phase I, but evidence indicates *no improvement* in the quality of health care;
3. The same contextual factors triggered *different mechanisms* than in Phase I, and evidence indicates *an improvement* in the quality of health care; or
4. The same contextual factors triggered *different mechanisms* than in Phase I, but evidence indicates *no improvement* in the quality of health care.

Only Outcome 1 would represent a full verification of the theoretical results from Phase I. If we see the same mechanisms triggered by the same contextual factors, we will

have more confidence that we have identified general tendencies of communities with certain characteristics to adopt certain approaches to T/A, as well as associated these mechanisms with improvements in health care (as in Phase I). Outcome 2 would represent a partial validation of the theory, in the sense that it correctly explained the role of context in determining the mechanism; but Outcome 2 would also suggest that the theory does not correctly specify all of the appropriate conditions for the mechanism to make a difference to the quality of health care. Thus Outcome 2 would call for additional evaluations by us and other scholars and practitioners, leading to further theoretical refinement, to understand what additional characteristics are necessary for the mechanisms to have an impact.

Outcomes 3 and 4 would represent a rejection of the Phase I theoretical insights, suggesting that they were at least partly biased by dynamics or elements unique to the two Phase I countries. That is, these outcomes would imply either that we did not correctly identify the characteristics that in a general sense (that is, across many developing countries) explain when a community will adopt a particular strategy to make use of the information a T/A intervention provides, or that there is no such general explanation for when a community will certain characteristics will adopt a particular strategy. Outcomes 3 and 4 would thus suggest at a minimum that much more theoretical work is needed, and may indeed suggest that a general theory of T/A's impact is unattainable. In this regard, Outcome 4 would provide less reason for optimism, as it is associated with no perceptions of health care improvement by any mechanism. Outcome 3 would suggest more reason for optimism than Outcome 4, because in Outcome 3 there

at least was some citizen perception of improvement in health care, even if it doesn't arise by the theorized mechanisms.

In analyzing the theoretical outcome from Phase II, whether 1, 2, 3, or 4, to design an appropriate strategy for follow-up research and practice, so as to further our goal of jumpstarting a virtuous interaction of research and practice adding systematically to the body of knowledge about whether, how, and in what conditions T/A can make a difference.

## V. CONCLUSION

In sum, this multi-country staged project will advance the field of transparency and accountability and contribute to the improvement of health services in developing countries in four distinctive ways. First, we will develop a new T/A intervention that builds upon prior work in this field, and which leverages the local knowledge of embedded and capable local partners through an intensive co-design process. Second, by conducting mixed-methods evaluations in several countries, we will determine the impact of this new intervention and begin to understand the reasons behind this impact, and the complex relationship between social conditions and mechanisms that transparency triggers. Third, the project will provide the basis for a nuanced and empirically verified theory of the impact of T/A on development and empowerment. Thereby we hope to answer the call of T/A scholars (McGee and Gaventa 2011, Joshi and Houtzager 2012, Glennerster 2005) for a theory to guide future research and evaluation of T/A—both the different mechanisms through which T/A interventions seek to affect development, such

as collaborative problem solving or community pressure, and the conditions under which these mechanisms should be expected to improve development outcomes and empower disenfranchised groups. Just as important, the result will be a range of new tools for practitioners of T/A as well as sectoral health experts wondering whether to add T/A interventions instead of or alongside more traditional health interventions.

## References

- Ackerman, John. 2004. "Co-Governance for Accountability: Beyond "Exit" and "Voice"." *World Development* no. 32 (3):447-463.
- Alsop, Ruth, and Nina Heinsohn. 2005. *Measuring Empowerment in Practice: Structuring Analysis and Framing Indicators*. World Bank Policy Research Working Paper 3510, World Bank, Washington, DC.
- Andrabi, Tahir, Jishnu Das, and Asim Khwaja. 2009. Report Cards: The Impact of Providing School and Child Test-scores on Educational Markets.
- Banerjee, A.V., E. Duflo, and R. Glennerster. 2008. "Putting a Band - Aid on a corpse: Incentives for nurses in the Indian public health care system." *Journal of the European Economic Association* no. 6 (2 - 3):487-500.
- Banerjee, Abhijit V, Rukmini Banerji, Esther Duflo, Rachel Glennerster, and Stuti Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* no. 2 (1):1-30. doi: 10.1257/pol.2.1.1.
- Bately, R. 1999. "The New Public Management in Developing Countries: Implications for Policy and Organizational Reform." *Journal of International Development* no. 11 (5):761-765.
- Beall, J., and M. Ngonyama. 2009. *Indigenous Institutions, Traditional Leaders and Elite Coalitions for Development: The case of Greater Durban, South Africa*, Leadership, Elites and Coalitions Research Programme, World Bank and Crisis States Research Centre, London.
- Benford, Robert D, and David A Snow. 2000. "Framing Processes and Social Movements: An Overview and Assessment." *Annual Review of Sociology* no. 26:611-639.
- Björkman, M., and J. Svensson. 2009. "Power to the people: evidence from a randomized field experiment on community-based monitoring in Uganda." *The Quarterly Journal of Economics* no. 124 (2):735-769.
- Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2005. "Missing in action: teacher and health worker absence in developing countries." *Journal of Economic Perspectives* no. 20 (1):91-116.
- Citizenship DRC. 2011. *Blurring the Boundaries: Citizen Action Across States and Societies*, Citizenship DRC, Institute of Development Studies, Brighton.
- Das, Jishnu, Jeffrey Hammer, and Kenneth Leonard. 2008. "The Quality of Medical Advice in Low-Income Countries." *Journal of Economic Perspectives* no. 22 (2):93-114.
- DFID. 2011. Empowerment and Accountability Summary Note. <http://www.dfid.gov.uk/Documents/publications1/empower-account-summary-note.pdf>.
- Duflo, Esther, Rema Hanna, and Stephen P Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* no. 102 (4):1241-1278. doi: 10.1257/aer.102.4.1241.
- Eyben, R., N. Kabeer, and A. Cornwall. *Conceptualising Empowerment and the Implications for Pro-Poor Growth*. Paper prepared for the DAC Poverty Network, Institute of Development Studies, Brighton.

- Fox, Jonathan. 2007. *Accountability politics: power and voice in rural Mexico*. New York: Oxford University Press.
- Fung, Archon. 2001. "Accountable Autonomy: Toward Empowered Deliberation in Chicago Schools and Policing." *Politics and Society* no. 29 (1):73-103.
- Fung, Archon, Mary Graham, and David Weil. 2007. *Full Disclosure: The Perils and Promise of Transparency*. New York: Cambridge University Press.
- Gaventa, John, and Gregory Barrett. 2011. "So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement." *IDS Working Papers* no. 2010 (347):01-72. doi: 10.1111/j.2040-0209.2010.00347\_2.x.
- Giugni, Marco G. 1998. "Was It Worth the Effort? The Outcomes and Consequences of Social Movements." *Annual Review of Sociology* no. 24:371-393.
- Glennerster, R. 2005. *Can information catalyze reform? Sierra Leone and rural India*, Abdul Latif Jameel Poverty Action Lab, MIT.
- Glewwe, Paul, and Eugenie W H Maïga. 2011. "The impacts of school management reforms in Madagascar: do the impacts vary by teacher type?" *Journal of Development Effectiveness* no. 3 (4):435-469. doi: 10.1080/19439342.2011.604729.
- Hastings, Justine S, and Jeffrey M Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments \*." *Quarterly Journal of Economics* no. 123 (4):1373-1414. doi: 10.1162/qjec.2008.123.4.1373.
- Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Holland, Jeremy, and Allyson Thirkell. 2009. *Measuring Change and Results in Social Accountability Work*. DFID Working Paper 34, DFID, London.
- Jenkins, Rob, and Anne Marie Goetz. 1999. "Accounts and Accountability: Theoretical Implications of the Right-to-Information Movement in India." *Third World Quarterly* no. 20 (3):603-622.
- Joshi, Anuradha. 2008. "Producing Social Accountability? The Impact of Service Delivery Reforms." *IDS Bulletin* no. 38 (6):10-17.
- Joshi, Anuradha, and Peter P Houtzager. 2012. "Widgets or Watchdogs?" *Public Management Review* no. 14 (2):145-162. doi: 10.1080/14719037.2012.657837.
- Keefer, Philip, and Stuti Khemani. 2011. "Mass Media and Public Services: The Effects of Radio Access on Public Education in Benin."
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing social inquiry: scientific inference in qualitative research*. Princeton, N.J.: Princeton University Press.
- Kosack, Stephen. 2012. *The Education of Nations: How the Political Organization of the Poor, Not Democracy, Led Governments to Invest in Mass Education*. New York: Oxford University Press.
- Lieberman, Evan S., Daniel N. Posner, and Lily Tsai. 2012. Does information lead to more active citizenship? In *An Evaluation of the Impact of the Uwezo Initiative in Kenya*.
- Mansuri, Ghazala, and Vijayendra Rao. 2013. *Localizing Development: Does Participation Work?* Washington, DC: World Bank.
- McAdam, Doug. 1982. *Political Process and the Development of Black Insurgency, 1930-1970*. Chicago: University of Chicago Press.

- McAdam, Doug, Sidney Tarrow, and Charles Tilly. 2001. *Dynamics of Contention*. New York: Cambridge University Press.
- McCarthy, John D, and Mayer N Zald. 1977. "Resource mobilization and social movements." *American journal of sociology* no. 86 (6):1212-1241.
- McGee, R., and J. Gaventa. 2011. "Shifting Power? Assessing the Impact of Transparency and Accountability Initiatives." *IDS Working Papers* no. 2011 (383):1-39.
- McGuire, James. 2010. *Health, Wealth, and Democracy in East Asia and Latin America*. New York: Cambridge University Press.
- Muralidharan, Karthik, and Venkatesh Sundararaman. 2010. "The Impact of Diagnostic Feedback to Teachers on Student Learning: Experimental Evidence from India\*." *The Economic Journal* no. 120 (546):F187-F203. doi: 10.1111/j.1468-0297.2010.02373.x.
- Narayan, D., and P. Petesch. 2007. "Agency, Opportunity Structure and Poverty Escapes." In *Moving Out of Poverty*, edited by D. Narayan and P. Petesch, 1-44. Washington, DC: World Bank.
- Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* no. 115 (2):200-249. doi: 10.1086/517935.
- Olson, Mancur. 1965. *The logic of collective action; public goods and the theory of groups*. Cambridge: Harvard University Press.
- Pandey, P., A.R. Sehgal, M. Riboud, D. Levine, and M. Goyal. 2007. "Informing resource-poor populations and the delivery of entitled health and social services in rural India." *JAMA: The Journal of the American Medical Association* no. 298 (16):1867-1875.
- Pandey, Priyanka, Sangeeta Goyal, and Venkatesh Sundararaman. 2009. "Community participation in public schools: impact of information campaigns in three Indian states." *Education Economics* no. 17 (3):355-375. doi: 10.1080/09645290903157484.
- Polletta, Francesca, and James M Jasper. 2001. "Collective Identity and Social Movements." *Annual Review of Sociology* no. 27:283-305.
- Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, and Rima Prama Artha. 2011a. Improving Educational Quality through Enhancing Community Participation. In *Results from a Randomised Field Experiment in Indonesia*.
- Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Arya Gaduh, and Rima Prama Artha. 2011b. Improving Educational Quality through Enhancing Community Participation: Results from a Randomised Field Experiment in Indonesia.
- Reinikka, Ritva, and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* no. 3 (2/3):259-267.
- Reinikka, Ritva, and Jakob Svensson. 2011. "The power of information in public services: Evidence from education in Uganda." *Journal of Public Economics* no. 95 (7-8):956-966. doi: 10.1016/j.jpubeco.2011.02.006.

- Ross, Michael. 2006. "Is Democracy Good for the Poor?" *American Journal of Political Science* no. 50 (4):860-874.
- Savage, J. 2000. "Ethnography and health care." *BMJ : British Medical Journal* no. 321 (7273):1400-1402.
- Stern, E., N. Stame, J. Mayne, K. Forss, R. Davies, and B. Befani. 2012. *Broadening the range of designs and methods for impact evaluations*. Working Paper 38, Research for Development, DFID, London.
- Tandon, R. 2003. *Civil Society and Policy Reforms*. IDS Civil Society and Governance Policy Brief No. 9, Institute of Development Studies, Brighton.
- World Bank. 2004. *World Development Report 2004: Making services work for poor people*. Vol. Oxford University Press: New York.
- Zweifel, T. D., and P. Navia. 2000. "Democracy, dictatorship, and infant mortality." *Journal of Democracy* no. 11 (99-114).



## APPENDIX: PHASE I SITE SELECTION CRITERIA

We use three criteria in selecting Tanzania and Indonesia for our Phase I countries:

1. *Service Provision Context (by our “five worlds” rubric):* We sought two sites within which we can find as many of the different service provision “worlds” identified in our rubric as possible;
2. *Baseline Contextual Factors:* We sought sites that vary on a number of potentially important contextual factors that, at baseline, may affect the likelihood of the success of a T/A intervention, the mechanisms by which it may have an effect, and the interaction of the intervention with power dynamics and relations. In particular, we sought variation on four sets of factors:
  - a. *Level of development*, which in part reflects the resources available for health care and which may also affect a range of other important aspects of the intervention, such as the level of existing infrastructure, urbanization, and perceptions of citizen efficacy;
  - b. *Features of the political economy of the health sector*, such as including considerations of how the degree of political competition may affect the willingness of policymakers and, indirectly, providers to engage in reform efforts, the role that service delivery plays in the legitimacy of the regime, and/or how decentralization and other reform efforts have affected accountability relationships between national and local politician, and policymakers/bureaucrats at multiple levels and providers; and

- c. *Culture, history, and geography*, which proxy for a wide range of factors that may be important to the effects of a T/A intervention;
3. *Strong Implementing Partners*: On this one criterion, we did not seek variation; rather we sought two sites that both have strong implementing partners that in particular are:
- a. *Learning organizations*, willing to innovate, co-design, and rigorously evaluate new approaches to T/A, and to accept both positive and negative results;
  - b. *Able to operate at the community level* and thus able to provide both the local knowledge that we argued above is crucial to designing an adaptable and locally appropriate intervention, as well as a platform for implementing the social action component of the intervention in ways that will take into account local variation and allow for adaptation and judgment;
  - c. *Operationally strong* and able to effectively execute interventions in a timely manner;
  - d. *Knowledgeable and experienced in the field of T/A* with extensive experience developing and implementing a wide variety of T/A interventions; and
  - e. *Interested in engaging with networks of other T/A stakeholders* and participating as a leader of thought and practice in the dissemination of the results of the evaluation.

We began with a shortlist of potential implementing partners focusing solely on criteria 3c and 3d above: CSOs with a strong track record of implementing high-quality and high-impact T/A work with the operational capacity to do so. This produced a preliminary list of approximately 30 potential partners across countries in Africa and Asia. We then conducted a series of in-person and phone meetings with potential partners to identify how well the organizations met other criteria.

This process led us to the selection of Tanzania and Indonesia as our Phase I sites. The two countries vary in their level of development (by the World Bank's categorization, Tanzania is a low-income country; Indonesia is a lower-middle-income country) and in myriad aspects of their culture, history, and geography. Indonesia is a democracy; by contrast, while multiple parties in Tanzania regularly contest elections, the country is functionally a stable, one-party-dominated state. Together, these two countries provide one setting, Tanzania, in which health policy will be highly consistent across communities (even if the implementation varies widely) and one, Indonesia, in which we will be able to exploit variation in health policy, and the willingness of policymakers to reform it, across communities. Finally the two have significant outreach and dissemination potential. Indonesia is playing an increasingly important role in T/A work in Southeast Asia: CSOs within the country are stepping up their work in this area, and more CSOs outside of the country are turning to Indonesian organizations for lessons learned. In fact, some Indonesian policymakers have also shown a commitment to incorporating T/A work into sectoral systems. Tanzania—in particular, one of the potential partners we are considering, Twaweza—has been a leader in innovating with new approaches to T/A interventions and many other CSOs within Tanzania have a more

regional reach across Africa in their project implementation and their dissemination, providing an audience base as well as partners to share wider findings during outreach.