International Public Management Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/upmj20

Accountability and effort among street-level bureaucrats: Evidence from a lab-in-the-field experiment

Alan Zarychta, Tara Grillos & Krister P. Andersson

To cite this article: Alan Zarychta, Tara Grillos & Krister P. Andersson (2021): Accountability and effort among street-level bureaucrats: Evidence from a lab-in-the-field experiment, International Public Management Journal, DOI: 10.1080/10967494.2021.1951408

To link to this article: https://doi.org/10.1080/10967494.2021.1951408

INTERNATIONAL PUBLIC MANAGEMENT JOURNAL
https://doi.org/10.1080/10967494.2021.1951408

Accountability and effort among street-level bureaucrats: Evidence from a lab-in-the-field experiment

Alan Zarychta[a], Tara Grillos[b], and Krister P. Andersson[c]

[a]University of Chicago; [b]Purdue University; [c]University of Colorado Boulder

ABSTRACT

Governance reforms like decentralization and performance-based management aim to improve public services by increasing accountability among street-level bureaucrats: bureaucrats may be held to account by communities, supervisors, intermediary organizations, or all of these. To assess the relationship between accountability and bureaucratic effort, we utilize data from a lab-in-the-field behavioral experiment conducted with Honduran health workers across decentralized and centrally administered municipalities. We presented health workers with an incentivized effort task that included instructions that were neutral, had a bottom-up political accountability prompt, or a top-down bureaucratic accountability prompt. Our results show that administrative context moderates the accountability-to-effort relationship. With neutral instructions, civil servants in decentralized systems exert greater quality effort than their counterparts under centralized administration. Importantly, both accountability prompts increase quality effort in centrally administered settings to levels comparable with those in decentralized settings. These findings support multiple accountability as a potentially important mechanism linking decentralization reform to improved service delivery.

ARTICLE HISTORY
Received 11 February 2020
Accepted 30 June 2021

Introduction

Creating and sustaining meaningful accountability is a central goal and predominant focus of contemporary public sector reforms (Bovens 2007; Brodkin 2007; Grindle 2004; Ospina, Grau, and Zaltsman 2004; Romzek 2000). The core idea of many of these reforms is that government performance will improve with greater accountability and thus members of the public will benefit from improved service delivery (Brodkin 2011; Hupe and Hill 2007). The emphasis on accountable *governance* is especially pronounced in efforts to strengthen local health systems in developing countries where performance-based incentives, contracting-out, and hybrid services delivery models coexist with, and sometimes take the place of, a traditional Weberian bureaucracy (Brinkerhoff 2004; Brinkerhoff and Bossert 2014; Ewert 2020; Siddiqi et al. 2009). In these types of increasingly common polycentric contexts where multiple principals and multiple agents jointly constitute the government performance experienced by the public, greater oversight and monitoring are often called for to counteract the problem of "muddled" accountability (Lieberman 2011; Meijer 2013). While it is expected by some and hoped by others that formal supervision by managers and informal vigilance by users will jointly create accountability, in practice a range of

CONTACT Alan Zarychta azarychta@uchicago.edu Crown Family School of Social Work, Policy, and Practice, University of Chicago, Chicago, IL 60637-5418, USA.

Supplemental data for this article is available online at https://doi.org/10.1080/10967494.2021.1951408.

individual, institutional, or contextual factors can prevent this from happening (Bauhr and Grimes 2014; Chong et al. 2015; Etzioni 2010; Etzioni 2014; Jayal 2007; Meier and O'Toole 2006).

In interrogating and assessing whether different types of monitoring can actually engender accountability, most research has focused on aggregate government performance or individual citizens' perceptions of the state, predominantly in Western settings. This work regularly acknowledges the importance of contextual conditions, but studies are often unable directly to assess those conditions (Cucciniello, Porumbescu, and Grimmelikhuijsen 2017). Additionally, as noted by de Boer and Eshuis (2018), few studies have connected notions of monitoring and oversight to the broader implementation literature that emphasizes the behavior of street-level bureaucrats as the critical agents that interface between the state and members of the public (Brodkin 2008; Hupe and Hill 2007; Lipsky 1980; May and Winter 2009). This last point, particularly, is a missed opportunity given that the expected system-level benefits of many accountability initiatives must necessarily be realized, at least in part, through changes in the behavior of street-level bureaucrats. Street-level bureaucrats are, in many ways, the lynchpin actors for contemporary public sector governance reforms (Andersson, Gibson, and Lehoucq 2006; Grillos, Zarychta, and Andersson 2021; Zarychta, Grillos, and Andersson 2020).

In this study, we draw on original data from a lab-in-the field framing experiment conducted with the participation of 127 health workers across 31 municipalities in Honduras with different administrative contexts. The experiment was implemented within a broader quasi-experimental study design based on an ongoing health sector decentralization reform in the country, and so some municipalities had decentralized administration of their health systems while others were centrally administered (Zarychta et al. 2019a). We engaged these street-level bureaucrats in a real effort task within which we embedded two different accountability prompts. In comparison to the control case where health workers received neutral instructions for the effort task, health workers in the treatment cases were prompted either that the community pays attention to the actions of health workers, the bottom-up political accountability frame, or that supervisors within the health system pay attention to the actions of health workers, the top-down bureaucratic accountability frame.

Our results show that administrative context moderates the effect of these accountability prompts on the quality of effort among doctors, nurses, and other health workers. Specifically, we find that under neutral instructions, health workers in the decentralized setting exert more quality-adjusted effort than their counterparts under centralized administration. The overall quantity of effort that civil servants exert, without considering quality, is not significantly affected by the accountability prompts or administrative contexts. Furthermore, our results point toward both bottom-up political and top-down bureaucratic accountability prompts increasing the quality-adjusted effort of centrally administered health workers up to the levels exhibited by their decentralized counterparts. This indicates that the possible benefits of decentralization may be driven in part by stronger expectations of monitoring and the greater salience of multiple accountability among health workers, both with respect to their communities and their supervisors.

This research makes four contributions to ongoing debates about the monitoring-accountability-performance relationship. First, we contribute a Latin American case to this literature where the conditions of public sector employment have important differences as compared to the United States and Europe (Finan, Olken, and Pande 2015). Second, we empirically assess the link between accountability prompts referencing two highly relevant types of monitoring that help shed light on the persistent debate over bottom-up, top-down, or multiple accountability in public services delivery (Ackerman 2004, Hill and Hupe 2006; Serra 2012). Third, we do so with actual street-level bureaucrats engaging in a real effort task, thereby advancing the type of behavioral-experimental approach that is becoming increasingly relevant to public management and administration (Battaglio et al. 2019; Carrigan, Pandey, and Van Ryzin 2020; James, Jilke, and Van Ryzin 2017). And fourth, our results provide empirical support for the importance of context

in understanding the effects of accountability (Brinkerhoff and Wetterberg 2016). Notably, we show evidence for how administrative form, centrally administered versus decentralized, is a key moderating condition on the relationship between accountability and effort among street-level bureaucrats.

Existing explainations of work effort by civil servants

Scholars have posited that the day-to-day work effort of civil servants may be influenced by individual characteristics, task attributes, organizational attributes, and the broader socio-cultural context (Franco, Bennett, and Kanfer 2002; Gailmard 2010; Moynihan and Pandey 2007; Pandey and Stazyk 2008; Wright 2004). In trying to understand the factors that drive some individuals to exert higher effort than others toward the public good, considerable research has focused on the idea of public service motivation, either as a necessary pre-requisite for effort, or effectively its equivalent. This is in part due to the prominent influence of the PSM perspective, one of the core hypotheses of which is that motivation will be positively correlated with performance on the job (Christensen, Paarlberg, and Perry 2017; Perry and Wise 1990).

The few experimental studies on this topic generally corroborate the original public service motivation hypothesis, though they often rely on survey-based measures derived from Perry's (1996) original, US-based scale (Bellé 2013; Pedersen 2015). The small number of studies employing behavioral measures also appear to support a positive relationship between motivation and worker effort (Ashraf, Bandiera, and Jack 2014; Banuri and Keefer 2016; Carpenter and Gong 2016; Oliveros and Schuster 2018). These studies generally address some of the limitations of survey measures, especially across different country contexts. However, several treat internal motivation and external effort as equivalent, even though these are distinct concepts with distinct behavioral measures.

Complementing this research on motivation, several scholars have also pointed out the importance of mission match in understanding the amount of effort that civil servants put into their work (Carpenter and Gong 2016; Resh, Marvel, and Wen 2018; Smith 2016; Wright 2007; Wright and Pandey 2011). This literature generally supports the notion that when individuals align with the missions of their organizations they will exert more effort in their work, and recent studies have focused on trying to identify the potential psychological mechanisms underpinning this relationship, including the perceived meaningfulness of the work and prosocial orientation (Resh et al. 2018; Smith 2016).

The notion of a compensating wage differential underlies much of the theoretical and empirical work across these studies on motivation, mission match, and effort: those individuals choosing to enter public service, with its commonly lower pay relative to the private sector, are necessarily thought to be motivated by the public good. Unlike many of the prominent cases in this literature, however, public sector employment in less developed country settings typically provides a wage premium over the private sector, job mobility can be quite limited, especially in rural areas, local institutions are relatively weak, and there are often acute concerns about corruption and the misuse of public resources (Finan et al. 2015; Ruhl 2010; Vandenabeele and Van de Walle 2008). Accordingly, while public service motivation and mission match can still be important, a greater emphasis on control and accountability is often appropriate in these cases, in line with research on implementation and state capacity.

Theoretical expectations on accountability and effort among street-level bureaucrats in developing countries

Accountability is often defined as the "… implicit or explicit expectation that one may be called on to justify one's beliefs, feelings, and actions to others" (Lerner and Tetlock 1999: 255). Despite

complexities and extensions, in the context of public administration and management the "core sense" of accountability is often considered to have four main features: it is external, it involves social interaction and exchange, it includes rights of authority, and it carries the possibility of sanctions (Mulgan 2000, 555). Building on this in a recent article, Han and Perry (2020) distill the concept of public employee accountability into five dimensions: attributability, observability, evaluability, answerability, and consequentiality. In short, a civil servant is accountable when they expect they will have to interact with and justify their actions to at least one other individual who has both a claim of authority over them and some ability to impose sanctions on them.

In a simplified version of a common decentralized administrative structure, there are two prominent groups that may be able to hold street-level bureaucrats or frontline service providers to account: (1) the clients, users, and other members of the public that they serve and who are situated below them at the end of the service delivery chain, and (2) the managers, administrators, and directors residing above them in the implementation hierarchy. Based on Romzek and Dubnick's (1987) classic presentation of four types of public accountability systems, the former corresponds to political accountability in that the degree of control over agency actions is relatively low and the source of that control is external to the agency, while the latter corresponds to bureaucratic accountability with its relatively higher control and its source being internal to the agency itself (229). In the context of decentralized governance reforms, we think it is important to further specify types of accountability as "bottom-up" or "top-down." This is because whether the relevant principal is situated above or below the agents in the service delivery hierarchy is an important conceptual distinction for assessing accountability in the context of reforms that aim to reallocate power and responsibilities across the levels of that hierarchy, as decentralization commonly aims to do (Dasandi and Esteve 2017). Following from a street-level approach to accountability (Brodkin 2008), we focus on these two types – bottom-up political accountability and top-down bureaucratic accountability – because they are the ones most directly experienced by frontline service providers and most commonly targeted for improvement as part of contemporary public sector governance reforms in developing countries (Brodkin 2007; Manning 2001; Ramos and Milanesi 2020).

Given this stylized version of a common hierarchical administrative structure, it is no surprise that principal-agent relationships, and their attendant problems, receive greatest attention in many studies of public sector reform in developing countries. The view from traditional, Weberian bureaucracy emphasizes political control and oversight of an apolitical civil service (Gailmard 2010). Control and oversight of the agent comes from a principal above them in the service delivery chain. The New Public Management, in contrast, recognizes decentralized models of administration and places priority on results. While there may be greater room for flexibility or discretion by agents under New Public Management, principals still play a major role in shaping their behavior, sometimes through market-based mechanisms and incentives (Hood 1991; Ramos and Milanesi 2020). Ossege (2012) describes this shift as one from process accountability to outcome accountability. Continuing this trend, some scholars have begun to emphasize a yet broader shift to *social* accountability, or what is sometimes called multiple accountability (Ewert 2020; Jayal 2007). The notion of social or multiple accountability attempts to go beyond the traditional single linkage models in order to consider the multiple relations and actors that may potentially be involved in holding bureaucrats to account, to allow for mechanisms related to voice alongside those centered on choice, and to include individual and collective values in understanding the behavior of civil servants (Hupe and Hill 2007; Jayal 2007; Meier and O'Toole 2006; Mulgan 2000; Schillemans, Van Twist, and Vanhommerig 2013).

This latter view, which most closely aligns with the reality of increasingly common polycentric governance arrangements for delivering social services, is one comprised of multiple principals and multiple agents (Brehm and Gates 2015). Theoretically, some scholars believe this can compound the standard principal-agent problem in terms of information asymmetries, attribution of

responsibility, and the construction of accountable governance (Gailmard 2009; Lieberman 2011). For example, Brinkerhoff (2004) argues that two few principal-agent linkages can be associated with corruption, too many may lead to muddled accountability, and internal linkages alone can privilege the interests of bureaucrats over the public. On the other hand, decentralization reforms and hybrid services delivery models may create opportunities for greater numbers and more diverse types of principals to hold bureaucrats to account, at least to some degree (Cuadrado-Ballesteros 2014; Escobar-Lemmon and Ross 2014; Ospina et al. 2004; Piatak, Mohr, and Leland 2017; Wright et al. 2016). In fact, recent research has highlighted cases in which combined top-down and bottom-up sources of accountability may be an important complement to state bureaucracy, or even substitutes for it, where traditional political institutions are weak (Cabral and Santos 2018; Cammett and MacLean 2011; Nelson-Nuñez 2019).

As a minimal condition, monitoring is critical for these complex varieties of accountability to be realized, and actually have the chance to change the behavior of the street-level bureaucrats who deliver social services (Ferry and Eckersley 2015; Meijer 2013). Monitoring refers to actions taken with the intention of observing, and if necessary, sanctioning the behavior of others (Gailmard and Patty 2012; Lancaster 2014). Monitoring is thus critical to accountability because diverse types of principals need to be able to find out what agents are doing, and those agents need to form credible expectations about the visibility of their actions and the potential that any associated sanctions would be applied. As Dubnick (2005: 399) argues, expectations are paramount within any accountability relationship, and without the preceding expectations in place, it is hard to imagine that social accountability would be able to engender more effort or better performance on the part of civil servants.

Our research therefore uses a framing experiment and real effort task to examine this link between civil servants' expectations of accountability and their work effort in the multi-principal setting of health sector governance in Honduras. The Honduran case is useful because the country's health sector decentralization reform combines two features that are prominent across contemporary strategies to improve accountability in service delivery: performance-based management that has a top-down character, and devolution that more closely reflects a bottom-up approach (Zarychta et al. 2019b). In line with calls to consider behavioral mechanisms and meso-level social contexts in studies of bureaucracy (e.g., Brehm and Gates 2015; Dubnick 2005), we assess the direct effects on effort of administrative form and top-down vs bottom-up accountability frames, as well as their interaction, to help shed light on the accountability-to-performance relationship.

With respect to the direct effect on effort of administrative form, decentralized versus centrally administered, most research has focused on policy outputs rather than effort directly, and this body of literature is notorious for its mixed findings (e.g., Abimbola, Baatiema, and Bigdeli 2019; Treisman 2007). More recent reviews, however, have concluded that decentralization can improve service delivery in health and education (Channa and Faguet 2016), and has increasingly come to focus on identifying mechanisms and contextual factors associated with improved performance under decentralization given wide variation in the details and implementation of these reforms (Abimbola et al. 2019). Our own work in this vein has shown that health sector decentralization in Honduras is credibly associated with increases in women's health services (Zarychta 2020) and broader improvements in maternal and child health (Root et al. 2020), both of which are types of production-based health indicators that require effort inputs from the doctors, nurses, and social workers within primary care health centers. Our work suggests these improvements have come about, at least in part, through the strengthening of accountability for frontline health workers under decentralization wherein those staff members have greater routine exposure to oversight and support than staff members in centrally administered health centers (Zarychta 2020). Accordingly, we hypothesize that effort will be higher on average among health workers in decentralized settings relative to their colleagues in centrally administered settings (H1).

In order to understand the consequences of bottom-up versus top-down accountability frames, we focus on the monitoring-to-effort literature. This literature has shown that both top-down monitoring interventions and interventions encouraging greater community-level monitoring of civil servants often fall short of their goals (Banerjee et al. 2010; Banerjee and Duflo 2006; Björkman and Svensson 2010; Dubnick 2005; Joaquin and Greitens 2011), though there is somewhat greater evidence in favor of top-down schemes (Dhaliwal and Hanna 2017; Olken 2007). Where there has been success, that success often hinges on different features of the local political context or features of the monitoring itself. For example, Véron et al. (2006) find that both upward and downward accountability are important in lowering the chances of corruption related to an employment assurance program in India, while Yanez-Pagans and Machicado-Salas (2014) show that monitoring by community-based actors can improve public service outcomes in Bolivia when other local institutions are relatively strong and free of elite capture. Focusing on management practices among civil servants in Nigeria, Rasul and Rogger (2018) find a negative relationship between monitoring and project completion rates, though that did not hold for project types with less complexity and greater certainty. Importantly, they also show that this negative correlation is offset by the proportion of intrinsically motivated individuals within an organization. In line with earlier work by Ostrom (2000) and Deci and Ryan (2000), this latter finding from Rasul and Rogger (2018) suggests that monitoring interventions might be able to crowd-in, rather than crowd-out, effort. Based on the increased prominence of oversight in general under the Honduran health sector decentralization reform, both top-down and bottom-up, and the relatively straightforward effort task we employ in our experiment, we expect that both accountability frames will be associated with higher effort simply by making monitoring more immediately salient to the civil servants participating (H2).

Following from considerable past research showing the importance of contextual factors in shaping the effectiveness of monitoring, we also hypothesize that administrative context will in fact moderate the accountability-to-effort relationship (H3). To the extent that performance incentives dominate community-based considerations under the decentralization reform, then the top-down accountability frame will be associated with greater effort in that setting (H3a). In contrast, if community-based considerations are more prominent than performance incentives under decentralization, then the bottom-up accountability frame will be associated with greater effort in that setting (H3b). Alternatively, if the possibility of broader social or multiple accountability holds in general, then one could expect that both of these accountability frames would be able to increase effort on this short-term task across the administrative settings (H3c). In assessing these relationships, we consider both quantity of effort and quality-adjusted effort in our experiment given prominent concerns about the abilities of officials to "game" public sector reforms, as well as specific concerns about the unintended consequences of performance incentives (Brodkin 2008; Heinrich and Marschke 2010). To the extent that concerns about narrowly defined and quantifiable performance have become inculcated through the reform and potentially affected what is viewed as valuable (Moynihan 2010), civil servants would express greater quantity (or raw effort) and lower quality effort. Absent these unintended consequences, one would not expect there to be differences between the quantity and quality of effort on a routine task.

Empirical strategy

To examine the relationship between monitoring, accountability, and effort, we conducted lab-in-the-field behavioral experiments with street-level bureaucrats in a matched sample of 31 municipalities in Honduras, some of which were managed under a decentralized health services delivery model and others which remained under the traditional, centrally administered model. In each municipality, we organized a daylong study-workshop during which actual health workers engaged in a series of activities and games, as well as a de-briefing discussion and capacity-

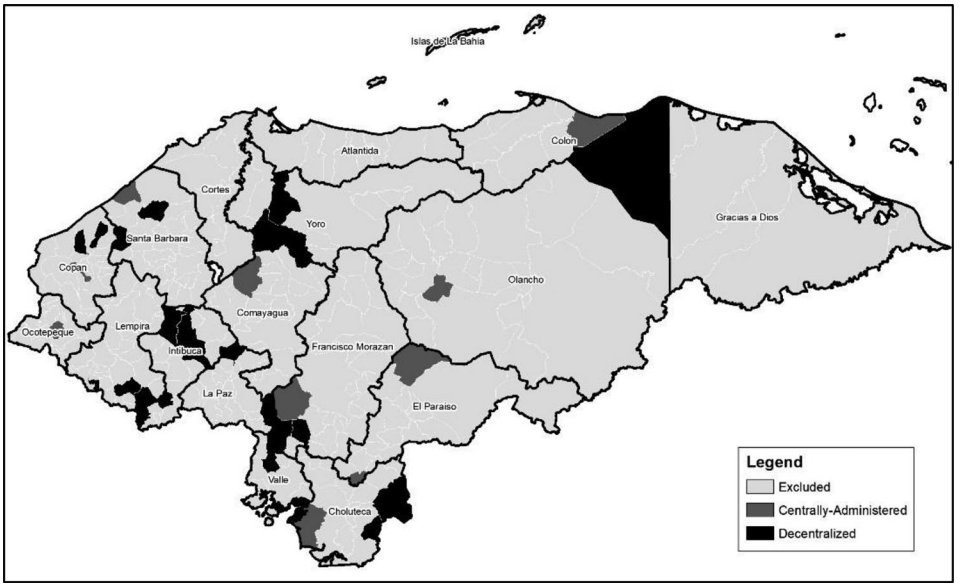

Figure 1. Honduran municipalities where lab-in-the-field experiments were conducted during workshop events in 2017.

building session related to health services administration. In the analysis presented here, we draw on data from the pre-survey and an effort task and associated framing experiment that constituted one of the day's activities. Additional details of the workshops are described in Zarychta et al. (2019b).

Sample of municipalities and health workers

The 31-municipality sample, shown in Figure 1, was chosen through a propensity score matching approach with respect to the ongoing health sector decentralization reform in Honduras. As part of a broader project and informed by interviews with key policymakers, we utilized pre-reform data on population, socio-demographic characteristics, health facilities, health services, distance, political factors, and other social interventions for all municipalities in Honduras to model the assignment of the decentralization reform following the iterative approach described in Imbens and Rubin (2015, 285–88). Based on the propensity scores from the resulting logistic regression model, nearest neighbor matching (with replacement, ratio 3:1) produced the most balanced sample among a set of commonly used matching strategies. Both the propensity score matching and the need to limit ourselves to no more than 3 municipalities given participation requirements for our workshops informed the selection of the sample of municipalities for this study. The final 31-municipality sample included 21 decentralized municipalities with an average propensity of decentralization of 0.64, and a matched set of 10 centrally administered municipalities where the corresponding propensity of decentralization was 0.58. Table 1 presents the overall and variable-by-variable balance assessment for this sample.

While this approach greatly improved balance between the two groups, to address the remaining minor differences in comparability we also created a reduced sub-sample where we dropped seven high-propensity decentralized municipalities that had less strong matches of centrally administered municipalities. In this sub-sample of 24 municipalities, 14 decentralized and 10 centrally administered, the two groups' propensities of decentralization were 0.47 and 0.49 respectively, and we use this sample for a set of robustness checks in Appendix B. The results of those robustness analyses are consistent with the finding presented here. Full details of all sample

Table 1. Centrally administered vs. Decentralized Balance for full sub-sample of 31 municipalities.

Variable	Centrally admin.	Decentralized	*P*-value diff.	SMD
Municipalities	10	21		
Propensity	0.58	0.64		
Total Primary Care HCs	6.59 (2.91)	6.00 (3.35)	0.651	0.187
Prop. Nurse-Only HCs	0.79 (0.10)	0.77 (0.12)	0.535	0.217
Pop. Per Primary Care HC	3,455.05 (1,204.73)	2,878.11 (1,169.59)	0.249	0.486
Total Population	24,426.50 (16,200.73)	16,859.33 (11,045.17)	0.257	0.546
Total Consults PP	2.03 (0.64)	2.23 (0.84)	0.479	0.260
First Prenatal Consults PP	1.14 (0.25)	1.06 (0.31)	0.458	0.289
Follow-up Prenatal Consults PP	3.02 (1.03)	2.68 (0.94)	0.356	0.336
Postpartum Consults PP	0.73 (0.30)	0.60 (0.21)	0.292	0.511
First Diarrhea Consults PP	0.18 (0.08)	0.22 (0.07)	0.211	0.565
Follow-up Diarrhea Consults PP	0.03 (0.02)	0.03 (0.03)	0.872	0.066
Growth Monitoring Consults PP	2.06 (0.67)	2.01 (0.85)	0.854	0.070
Distance to Tegucigalpa	87.32 (70.95)	118.39 (57.53)	0.230	0.481
Distance to State Capital	22.88 (14.53)	29.33 (12.96)	0.261	0.469
Prop. Indigenous Pop.	0.05 (0.18)	0.19 (0.21)	0.031	0.712
Civil Society Orgs.	2.27 (2.53)	1.57 (2.04)	0.511	0.304
Cash Transfer Beneficiaries PP	5.06 (4.59)	3.62 (4.18)	0.444	0.327
Participation Mayoral Election	0.61 (0.10)	0.58 (0.11)	0.496	0.262
Margin Victory Mayoral Election	0.19 (0.13)	0.12 (0.09)	0.189	0.596
Prop. Null Votes	0.09 (0.04)	0.10 (0.04)	0.669	0.174
Liberal Party Mayor	0.35 (0.50)	0.52 (0.51)	0.424	0.344
National Party Mayor	0.65 (0.50)	0.48 (0.51)	0.424	0.344
Libre Party Mayor	0.00 (0.00)	0.00 (0.00)	NA	<0.001
Other Party Mayor	0.00 (0.00)	0.00 (0.00)	NA	<0.001
Life Expectancy	69.91 (0.78)	69.45 (0.87)	0.162	0.553
Human Development Index	0.59 (0.02)	0.57 (0.04)	0.066	0.634
Education Index	0.58 (0.05)	0.55 (0.08)	0.112	0.573
Income Index	0.48 (0.02)	0.46 (0.04)	0.180	0.447
Health Index	0.75 (0.01)	0.74 (0.01)	0.165	0.551
Est. Income	1,743.93 (213.76)	1,633.14 (442.55)	0.337	0.319
Literacy Rate	70.68 (4.54)	66.13 (9.05)	0.067	0.636
Schooling Rate	33.93 (6.82)	31.44 (6.57)	0.348	0.371
Fiscal Autonomy Index	36.10 (20.74)	23.15 (17.56)	0.144	0.674

Notes: Table re-produced from Zarychta et al. (2019b); weighted averages by group with standard deviations in parentheses.

selection procedures are available in Zarychta et al. (2019b). We include municipal-level weights based on the relevant propensity scores in all of the subsequent analyses to ensure that decentralized and centrally administered municipalities are as comparable as possible.

Within each of the 31 selected municipalities, we received support from the Honduran Ministry of Health (MOH) to recruit a systematic sample of public sector doctors, nurses, social workers, health administrators, and municipal officials to participate in our workshops. Specifically, officials from the central MOH and from the relevant regional health authorities provided us lists of designated staff members for the municipalities in our sample representing the major roles within each municipal health system for us to invite to the workshop. The exact composition of individuals differed slightly based on the location, but our typical list included the following: doctors, nurses, and social workers from the municipality's primary care health centers (e.g., health workers), administrators from the regional health authority, administrators from the decentralized managing organization (if decentralized), and a health representative from the municipal government. Having the list of designated participants defined, we then sent formal letters through the MOH to request their participation; in the centrally administered cases, the Undersecretary of Health was the signatory on the letter, and in the decentralized cases, this was the Director of the Unit for Decentralized Management. Finally, our field staff conducted extensive follow-up via phone and in-person to facilitate participation by the designated individuals or to identify appropriate replacements given the role a particular individual represented within the health system. We made every effort to ensure that the individuals designated and recruited to

Table 2. Descriptive statistics for health workers sample.

	N	Min	Median	Mean	Max	SD	IQR	#NA
Total letters	126	7	14.50	14.84	25	3.62	5.00	1
Properly completed letters	126	0	12.00	11.33	25	5.07	5.75	1
Prop. properly completed letters	126	0	0.86	0.76	1	0.28	0.37	1
Age	127	21	31.00	34.10	55	8.72	13.00	0
Female	127	0	1.00	0.69	1	0.47	1.00	0
Years working health	127	0	5.00	7.43	30	6.91	10.00	0
Years current job	127	0	3.00	5.83	30	6.66	8.00	0
Education	124	2	5.00	4.46	5	0.69	1.00	3
Cooperative	127	1	1.00	1.04	2	0.20	0.00	0
Generalized trust	127	0	0.00	0.43	1	0.50	1.00	0
Particular trust	127	1	3.00	3.24	5	0.87	1.00	0
Doctor	127	0	1.00	0.52	1	0.50	1.00	0
Nurse	127	0	0.00	0.43	1	0.50	1.00	0
Social worker	127	0	0.00	0.04	1	0.20	0.00	0
Health knowledge quiz	122	0	0.60	0.58	1	0.22	0.30	5

participate based on their roles in the health system were able to attend, including re-scheduling workshops and providing transportation if needed, and there were only a very few cases where designated individuals were unable to make it to the events. Overall, a systematic sample of 232 civil servants representing the critical roles in the country's municipal-level public health systems participated in our workshops. Given the goals of the framing experiment described in the next sections, we focus this analysis on the sub-group of 127 health workers, excluding administrators and municipal officials. Descriptive statistics for this sample of health workers are provided in Table 2.

Experimental design: Accountability prompts and effort task

During the course of each workshop, health workers had the opportunity to participate in an effort-for-charity task. This was a framed version of the standard envelope stuffing task that is widely used for measuring real effort because it is simple, requires no previous knowledge, and is identical across repetitions (Carpenter and Gong 2016; DellaVigna et al. 2016; Falk and Ichino 2006; Gill and Prowse 2011; Konow 2000).

Specifically, we gave participants the option to assist us in preparing letters that we would send to relevant actors in the health sector informing them about the preliminary findings from our prior research in Honduras. Through their voluntary participation in this activity, they could earn donations for their chosen health centers. The activity included a basic set of instructions that were to fold each letter in thirds, insert it into the envelope, seal the envelope, and add stickers for the mailing address and return address in the correct places on the envelope. These instructions were common to all participants and meant to be benign with respect to effort. They were, however, intended to reflect the kinds of details common to administrative tasks within health centers, for example coding patient encounters or completing required forms that tabulate health services provided over the month. The purpose of this was to guard against, or at least be able to observe, the kind of opportunism that is sometimes associated with performance incentives (e.g., a situation where the envelopes are carelessly stuffed just to maximize the incentive pay but without attention to the underlying goal of activity which is to produce letters with the proper information so that they will actually reach their intended recipients). Participants, if they voluntarily chose to engage in this activity, had 15 minutes to prepare as many envelopes as they could. For each properly completed envelope at the end of the time, we agreed to donate five Honduran lempiras to a local health center of their choice.

In this case, since the envelope stuffing task requires effort but gives no personal benefit to the individual, it provides an excellent behavioral measure of the participant's willingness to engage

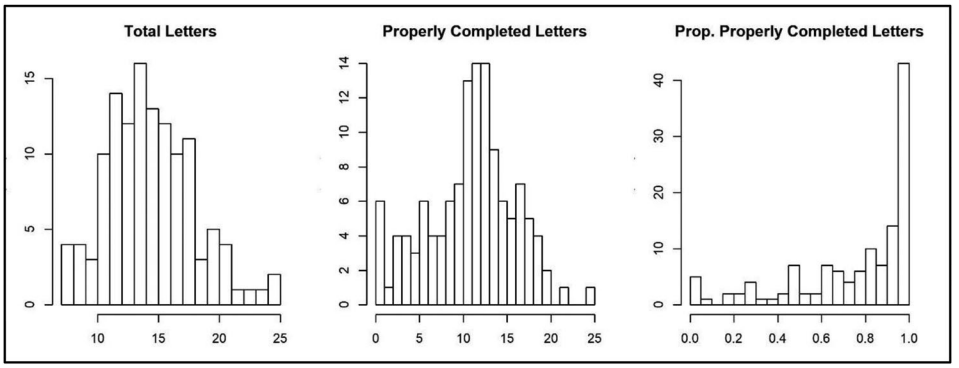

Figure 2. Measures of effort for health workers.

in effort in order to further the goals of the health sector. We recorded both the total number of envelopes completed and the number completed properly in order to differentiate between quantity effort, namely their total output of any kind during the activity, and quality effort, as reflected by paying attention to and following the basic instructions of the task. Figure 2 presents histograms of both effort measures for the full sample of health workers.

We incorporated a framing experiment for health workers within the effort task to assess the direct and conditional influence of different types of accountability prompts in this context. By "framing experiment," we mean an experiment in which the randomized manipulation involves changing how things are described, or "framed," rather than changing the tasks, choice set, or material consequences to which experimental subjects are exposed. Tversky and Kahneman's (1981) landmark study demonstrated that simply describing alternatives in terms of losses as opposed to gains leads people to make more risk-taking decisions. More recently, framing experiments have become popular in political science to establish the general phenomenon whereby "(often small) changes in the presentation of an issue or an event produce (sometimes large) changes of opinion" (Chong and Druckman 2007, 104). These "framing effects" have been explored with respect to self-stated opinions on issues ranging from welfare assistance (Rasinski 1989) to allowing hate groups to hold a rally (Sniderman and Theriault 2004), but in contrast to the work we present here, these related studies typically take the form of survey experiments where no actual behavior is directly observed. As Chong and Druckman (2007) note, the effects studied by framing experiments are also closely related to priming, in that presenting a different frame primes subjects to consider different aspects of an issue. In our case, for example, we prime subjects to consider different forms of accountability as they engage in our effort task.

A related methodological trend has seen the rise of lab-in-the-field experiments (Grossman 2011), in which researchers apply elements of conventional lab experiments "but take advantage of a particular naturally occurring situation in the field" (Morton and Williams 2010, 296). This is closely related to the idea of a "framed field experiment," which is also defined in contrast to traditional laboratory experiments by their use of both subject pools and contexts directly related to the field setting under study (Harrison and List 2004, 1014). Our study represents a combination of these two methodological approaches: framing experiments and lab-in-the-field experiments. We implement conventional laboratory techniques (effort tasks) in a field setting (the Honduran health sector), and we experimentally manipulate the way in which the task is framed with respect to accountability. Participants in lab-in-the-field experiments likely bring their prior experiences and perceptions of each other and of the context into the activity (Cárdenas and Ostrom 2004), and the effectiveness of framing has been shown to depend on context-specific preexisting beliefs (Andrews et al. 2013). In our setting, then, subjects' perceived likelihood of

Table 3. Sample of health workers by accountability frame treatment and administrative form.

Administrative form (blocking)	Accountability frame			
	Neutral (control)	Bottom-up (community)	Top-down (supervisor)	*Total*
Centrally administered	14	14	13	41
Decentralized	28	29	29	86
Total	42	43	42	127

monitoring by different agents should affect the degree to which they respond to different accountability frames. As such, the strategy of combining a framing experiment with a framed field experiment is well justified for our research question.

In our lab-in-the-field framing experiment, we developed three different versions of the instructions that participants received at the beginning of the effort task (included in Appendix A). One set of instructions detailed the basic features of the effort task as described above with no additional information (neutral control framing).

The two treatment versions of the instructions included an additional sentence at the beginning of the second paragraph that provided an accountability frame to participants. This sentence prompted the participant to recall that certain actors pay particular attention to efforts made by staff in raising additional resources to improve the services provided by the local health centers. In the *bottom-up political accountability* treatment, this actor is the community: "As you know, in many cases **the Community** pays attention to the efforts of people working to raise additional resources to improve the services provided by the municipality's health centers" [translated from Spanish, boldface in original]. In the *top-down bureaucratic accountability* treatment, the boldface text in the sentence references the unit within the Ministry of Health that is tasked with monitoring health center workers. This unit has a different title between the centrally administered and decentralized municipalities, but in both cases we used the reference for the supervisor immediately above the health workers in the formal hierarchy: the **Network Team of the Regional Health Authority** and the **Technical Team of the Decentralized Managing Organization**, respectively. Only health workers received one or the other of the treatment instructions because they are direct producers of health services located between the community and the supervising administrators, so these accountability frames would be of greatest applicability to them. Table 3 summarizes our 127-health worker sample by treatment category and administrative form.

Finally, Table 4 presents a balance analysis for participants in our framing experiment using standardized mean differences (SMD) across the neutral control, bottom-up political accountability, and top-down bureaucratic accountability instructions. We note some signs of imbalance in terms of years working in the health sector and years in current job, with somewhat longer-serving health workers being in the neutral control category. We verified that our randomization procedures were sound, and so we attribute this to the somewhat limited sample sizes per category. All remaining attributes appear well balanced, and we include regression models with a full set of individual controls in the subsequent analyses.

Analysis and results

Following from the experimental design described in the previous section, our analysis considers the effects of each accountability prompt (bottom-up political or top-down bureaucratic) on both quantity and quality of worker effort as moderated by the administrative form of health system (decentralized or centrally administered). Acknowledging sample size limitations, there are two principal conclusions that emerge from the subsequent analysis. First, under neutral conditions without any accountability frame, health workers experiencing decentralization exert greater *quality* effort than their counterparts in comparable centrally administered health systems, which

Table 4. Sample balance by accountability treatments.

	Neutral (control)	Bottom-up accountability (community)	Top-down accountability (supervisor)	SMD
Age	35.26 (9.01)	33.29 (8.09)	33.96 (8.74)	0.151
Female	0.68 (0.47)	0.70 (0.46)	0.67 (0.48)	0.045
Years working health	9.36 (8.29)	6.53 (6.01)	7.27 (6.93)	0.259
Years current job	8.31 (8.27)	3.86 (4.33)	5.48 (6.66)	0.446
Education	4.41 (0.78)	4.53 (0.62)	4.47 (0.65)	0.121
Cooperative	1.04 (0.20)	1.03 (0.18)	1.07 (0.25)	0.110
Generalized trust	0.51 (0.51)	0.35 (0.48)	0.40 (0.50)	0.205
Particular trust	3.15 (0.85)	3.47 (0.91)	3.15 (0.76)	0.248
Doctor	0.51 (0.51)	0.49 (0.51)	0.56 (0.50)	0.095
Nurse	0.44 (0.50)	0.47 (0.51)	0.39 (0.49)	0.103
Social worker	0.05 (0.22)	0.02 (0.15)	0.05 (0.22)	0.090
Health knowledge quiz	0.56 (0.23)	0.57 (0.19)	0.61 (0.22)	0.152

Notes: Weighted means using full health workers sample.

Table 5. Average effort outcomes for health workers across accountability treatments.

	Total letters	Properly completed letters	Proportion properly completed letters
Neutral framing (control)	15.38	11.28	0.73
Bottom-up accountability (community)	14.77	11.64	0.78
Top-down accountability (supervisor)	14.70	12.04	0.82

Notes. Weighted means using full health workers sample.

supports our first hypothesis. There are no significant differences in quantity effort, but trends suggest that increases in quality effort come from decreases in total output, at least in part.

Second, in line with the general orientation of our third hypothesis, administrative form is an important moderator of the relationship between accountability and effort. Specifically, both accountability frames appear to influence health workers in centrally administered municipalities to increase their quality effort up to the already higher levels exhibited by health workers in the decentralized municipalities. This is consistent with our second hypothesis, but only appears to hold for health workers in the centrally administered settings. Finally, considering hypotheses 3a, 3b, and 3c, our results suggest that accountability is already a salient expectation among health workers under decentralization relative to their centrally administered counterparts, and that this expectation is neither accentuated nor diminished by either accountability prompt. For the centrally administered health workers, accountability appears to be less salient in general, but both bottom-up political and top-down bureaucratic accountability prompts serve equally to increase their quality effort, at least in the short-term.

Bottom-up political versus top-down bureaucratic accountability and effort

Table 5 below presents average outcomes on our quantity and quality effort measures for the three experimental categories. Both accountability frames, bottom-up political and top-down bureaucratic, show slightly more properly completed letters in the effort task, in absolute terms and as a proportion of total letters. Notably, we see the opposite descriptive trend in terms of total letters; those health workers receiving the accountability prompts produce slightly fewer total letters during the effort task, but they have a higher proportion of their letters completed properly as described in the task instructions.

Table 6 presents a set of propensity score weighted OLS regression models of effort by health workers during our experiment as a function of the accountability treatments and a set of individual and municipal controls. We carry through the distinction between total and properly completed letters in this analysis, with the models for properly completed letters including a control for total letters, and presenting base and full models for both outcomes. In the full models we

also adjust for female versus male, education, age, years of experience working in the health sector, whether there was a foreign team member present at the workshop, and whether the municipality was also part of the Salud Mesoamerica Initiative (Mokdad et al. 2015). Considering the full model for properly completed letters, we see a slight increase in effort for both the bottom-

Table 6. Explaining effort outcomes for health workers by accountability frames.

	Total letters, base	Total letters, full	Proper letters, base	Proper letters, full
Bottom-up (Ref: Neutral control)	−0.61	−0.71	0.88	0.80
	(0.82)	(0.87)	(0.92)	(0.85)
Top-down (Ref: Neutral control)	−0.68	−0.74	1.34	1.21
	(0.77)	(0.86)	(1.04)	(1.02)
Female		1.29**		−0.41
		(0.64)		(0.76)
Education		1.45**		1.34*
		(0.57)		(0.68)
Age		−0.12		0.01
		(0.08)		(0.08)
Years working in health		0.02		−0.02
		(0.09)		(0.08)
Observer present		0.40		0.48
		(0.74)		(1.11)
Salud Mesoamerica		0.78		0.22
		(0.70)		(0.92)
Total letters			0.86***	0.82***
			(0.14)	(0.16)
Constant	15.38***	11.71***	−1.92	−7.30*
	(0.66)	(2.58)	(2.16)	(4.33)
R^2	0.01	0.19	0.38	0.44
Num. obs.	126	123	126	123

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$; health workers sample, weighted 31 munis., SEs clustered by muni.

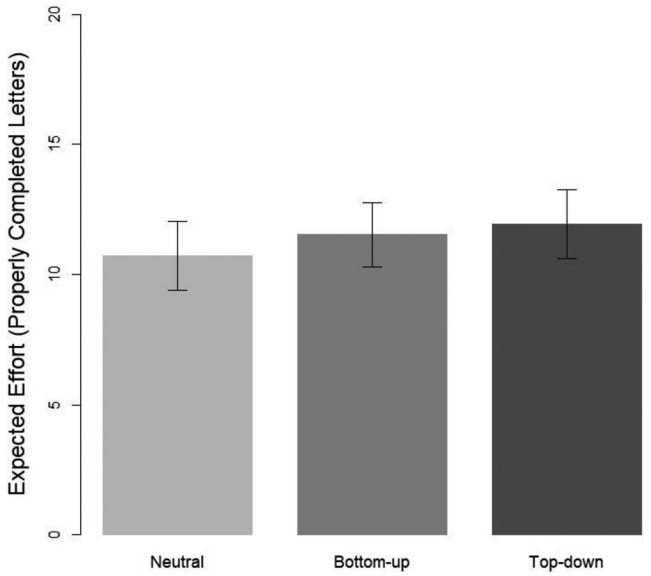

Figure 3. Expected effort for a typical health worker by accountability framing treatment (based on "proper letters, full" model from Table 6).

up and top-down accountability prompts for our sample of health workers, though neither is statistically significant at conventional levels.

Figure 3 below displays this positive but insignificant association between the accountability prompts and effort for a typical health worker in our sample using a simulation-based approach and showing 90-percent confidence intervals (Carsey and Harden 2013). In the neutral control category, the typical health worker is expected to complete about 10.7 proper letters during the effort task. With the bottom-up political accountability frame this rises to 11.5 letters and to 11.9 letters with the top-down bureaucratic accountability frame. These increases reflect about 7 and 10% of the sample average (11.3 properly completed letters), but counter to H2, we cannot conclude that the accountability prompts had a strong and unconditional influence on effort across the whole sample of health workers.

Administrative form and the accountability-to-effort relationship

Table 7 below presents average outcomes on our quantity and quality effort measures for the three experimental categories by the two administrative forms. Most notably, we see that health workers in centrally administered municipalities who received the neutral control instructions completed about one fewer letter in total and more than three fewer proper letters as compared to their counterparts in decentralized municipalities who also received neutral instructions for the task. For quantity effort in terms of total letters, the descriptive data suggest fairly similar effort for both accountability prompts in both administrative contexts, with a minor increase in raw output among decentralized staff receiving neutral instructions. In terms of properly completed letters, we see a slight increase in effort with the top-down accountability prompt in decentralized settings as compared to the centrally administered context, while the bottom-up accountability prompt produced similar numbers for properly completed letters across the two administrative settings. Within the centrally administered setting, both accountability frames are associated with an increase of more than two properly completed letters, but there do not appear to be differential effects between the two frames. Together, these descriptive results point toward both accountability prompts increasing quality effort among the health workers in centrally administered settings up to levels similar to those for staff in the decentralized setting, but not influencing the overall quantity of effort exerted, nor further or differentially affecting effort for staff managed under the decentralized model.

Table 8 presents the same set of propensity score weighted OLS regression models of effort as before, simply replacing the treatment dummy variables with a set of dummy variables that interact the accountability treatments with administrative form. For both the base and full models of properly completed letters, we do see evidence of statistically significant increases in effort for the treatments relative to the centrally administered-neutral framing category.

Figure 4 below displays the expected number of properly completed letters for a typical health worker in our sample using a simulation-based approach as in the prior analysis, in this case based on the full, proper letters model in Table 8. Comparing the first and fourth bars in

Table 7. Average effort outcomes for health workers across accountability treatments blocked by administrative form.

Administrative form	accountability framing	Total letters	Properly completed letters	Proportion properly completed letters
Centrally administered	Neutral	14.64	8.84	0.61
Centrally administered	Bottom-up	14.98	11.59	0.77
Centrally administered	Top-down	14.91	11.41	0.78
Decentralized	Neutral	15.69	12.31	0.77
Decentralized	Bottom-up	14.66	11.67	0.78
Decentralized	Top-down	14.62	12.31	0.84

Notes: Weighted means using full health workers sample.

Table 8. Explaining effort outcomes for health workers by accountability frames and administrative form.

	Total letters, base	Total letters, full	Proper letters, base	Proper letters, full
Decent, neutral (Ref: Central, neutral)	1.05	0.12	2.58**	2.49
	(1.06)	(1.43)	(1.29)	(1.55)
Central, bottom-up (Ref: Central, neutral)	0.35	−0.46	2.46	3.21**
	(1.39)	(1.32)	(2.16)	(1.48)
Decent, bottom-up (Ref: Central, neutral)	0.02	−0.73	2.81**	2.41
	(0.87)	(1.10)	(1.27)	(1.57)
Central, top-down (Ref: Central, neutral)	0.27	0.46	2.34	2.78
	(0.75)	(0.96)	(2.15)	(1.97)
Decent, top-down (Ref: Central, neutral)	−0.02	−1.13	3.49***	3.17**
	(0.90)	(1.10)	(1.28)	(1.41)
Female		1.26*		−0.38
		(0.67)		(0.82)
Education		1.56***		1.40**
		(0.59)		(0.69)
Age		−0.12		−0.02
		(0.08)		(0.08)
Years working in health		0.01		0.05
		(0.08)		(0.07)
Observer present		0.38		0.73
		(0.79)		(1.13)
Salud Mesoamerica		0.96		−0.13
		(0.85)		(1.05)
Total letters			0.85***	0.81***
			(0.14)	(0.16)
Constant	14.64***	11.30***	−3.56	−8.77**
	(0.56)	(2.72)	(2.36)	(4.06)
R^2	0.01	0.21	0.40	0.45
Num. obs.	126	123	126	123

Notes: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$; health workers sample, weighted 31 munis., SEs clustered by muni.

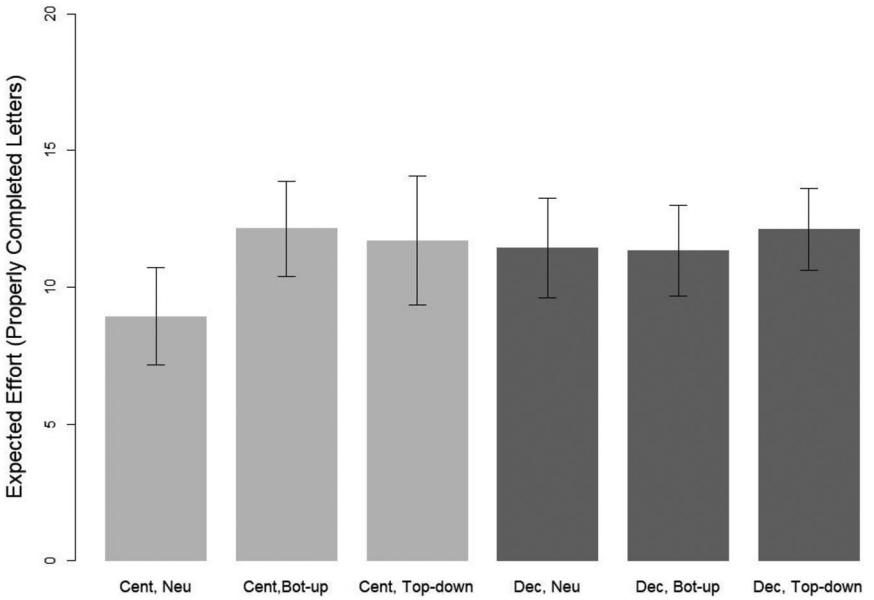

Figure 4. Expected effort for a typical health worker by accountability framing treatment and admin. Form (based on "proper letters, full" model from Table 8).

Figure 4, there is statistical confirmation of the positive association between decentralization and quality-adjusted effort discussed above: with no reminder of accountability, a typical health worker from our sample properly completed about 3 more letters under decentralization than

that same health worker would have been expected to complete in a centrally administered health system. Furthermore, the figure shows that being presented with an accountability frame, either top-down bureaucratic or bottom-up political, appears to increase the quality-adjusted effort of health workers in the centrally administered setting up to the already higher levels exhibited by health workers under decentralization: bars two and three in the figure are not statistically different from bars three, four, and five. Given limitations in sample size, we consider these findings to be suggestive and worthy of additional assessment. Nonetheless, they do indicate 2-3 additional properly completed letters over a 15 minute time period were associated with the accountability prompts and decentralization reform, which reflects a meaningful increase of 17- to 26-percent relative to overall sample average of 11.3 properly completed letters.

Discussion and conclusions

This study assessed the relationship between accountability and effort using data from a lab-in-the-field behavioral experiment conducted with Honduran health workers. We considered accountability in terms of two types of monitoring that are prominent across contemporary service delivery reforms: reminding health workers in the instructions to an effort task that either the community (bottom-up political accountability frame) or their supervisors (top-down bureaucratic accountability frame) pay attention to their actions. By utilizing this framed and incentivized real effort task with street-level bureaucrats, we also highlighted an important distinction between external effort and internal motivation, two concepts that are often used interchangeably.[1] Finally, we implemented our experiment in the context of an ongoing health sector decentralization reform, which also allowed us to assess the potential moderating role of administrative context on the accountability-to-effort relationship. In doing so, this work extends our understanding of the drivers of effort among public sector workers in developing countries, and how effort can be shaped by accountability prompts and institutional reforms in order to improve social services.

The results of our experiment showed that there is higher quality effort among health workers in decentralized settings, and that reminding health workers in the centrally administered settings of either bottom-up political or top-down bureaucratic accountability increased their quality-adjusted effort up to levels comparable with those of health workers in decentralized systems. Additionally, the overall quantity of effort exerted by health workers on this task, without considering quality, was not influenced significantly by differences in the accountability treatments or administrative contexts. These findings suggest that accountability and oversight are already salient expectations for health workers experiencing decentralization, and is evidence in favor of it being possible to engender some degree of multiple or social accountability within complex governance arrangements (Cuadrado-Ballesteros 2014; Escobar-Lemmon and Ross 2014; Ewert 2020; Ospina et al. 2004). This finding is in line with other research on health sector reform in Honduras showing that accountability was strengthened for frontline health workers under decentralization through greater numbers of monitoring visits at the health center level, and that this played a role in increasing the production of health services (Zarychta 2020). It is also some evidence against the notion that performance-based accountability, or other top-down variants, will necessarily have negative consequences for the behavior of civil servants, at least in terms of quality effort on routine tasks (Ossege 2012). In short, expectations of monitoring and accountability are material to the effort exerted by street-level bureaucrats, and there is evidence that those expectations can be shifted positively through both administrative reform and accountability prompts.

Additionally, this work speaks to the ongoing debates over bottom-up, top-down, or multiple accountability in improving service delivery, especially in the context of contemporary public sector governance reforms that include aspects of decentralization alongside aspects of performance-

based management. At least among the actual civil servants in this case, it appears that a combined approach to accountability can make both communities and supervisors salient principals and thereby help increase the quality of effort exerted in common work tasks (Ackerman 2004; Serra 2012). This remained true even in the context of a decentralization reform that includes aspects of performance-based management with incentives tied to results. Health workers interact directly with their communities and patients on a day-to-day basis, and top-down bureaucratic accountability with its internal character and relative strength did not crowd-out or override the bottom-up political accountability linkage as some might have feared. For Honduran health workers managed under the decentralized service delivery model, their top-down principals are more proximate to them as compared to health workers managed under the centralized model. This greater proximity is one of the key structural changes of the reform, and it is possible that it facilitated greater alignment on outcomes and accountability between the agents and their principals. This condition of better-aligned expectations, which requires additional investigation to confirm, could help explain why there was not necessarily a tradeoff between top-down versus bottom-up accountability, and that concerns about muddled responsibility or competing principals did not appear to have detrimental consequences in this case (Lieberman 2011).

Furthermore, the findings of this research highlight the importance of going beyond the standard principal-agent framework to consider how administrative and institutional contexts shape monitoring, accountability, and effort. The accountability framing treatments utilized in this study, which echoed common types of monitoring interventions for improving bureaucratic accountability in developing countries, had their most pronounced effects only when taking account of differences in administrative context. Our findings about the moderating role of administrative context also align with other recent work emphasizing the importance of state structures, processes, and state-society linkages, notably the degree of decentralization, in creating the conditions necessary for social accountability (Brinkerhoff and Wetterberg 2016). Nonetheless, as others have suggested, this remains an area that would benefit from additional observational and experimental research (Cucciniello et al. 2017; Meier and O'Toole 2006).

While our study design and implementation choices were made in an effort to improve on existing work, the findings of this research should be interpreted in light of certain limitations. First, collaborating with the Honduran MOH and engaging real health workers in this research placed limits on the sample sizes we could achieve, as did resource constraints, which qualify some of the conclusions we reach. Second, our study draws on evidence from a framing experiment conducted within a quasi-experimental design where centrally administered and decentralized municipalities were matched using propensity scores. In our view this was a strong and credible research design in this context given the barriers to randomizing the targeting and implementation of governance reforms like decentralization (Zarychta et al. 2019a), but nonetheless matching does have limitations as a strategy for causal inference, which appropriately tempers the interpretation of the study's findings. Third, the accountability framings we incorporated into the instructions for the effort task are one-time prompts deployed in the context of a workshop, and follow-up studies would be well served to consider stronger and more frequent prompts that might yet better approximate how bureaucrats experience monitoring day-to-day. Relatedly, we focused this analysis on service providers, namely doctors, nurses, and social workers who regularly interface with community members, but additional research is needed to better understand the accountability-to-effort relationship for managers and administrators above those providers within the service delivery system. Lastly, while our effort task mirrors a set of common activities integral to health workers' day-to-day work, it is nonetheless an immediate and short-run measure of effort. As such, future research is needed to assess what factors are important for the persistence of increased effort among street-level bureaucrats.

Despite these limitations, our research helps scholars, policymakers, and managers consider the appropriateness of general accountability reforms, as well as the design of specific monitoring policies. Moreover, it does suggest that there are possibilities for social accountability in governance settings with multiple principals and multiple agents. As other researchers have begun to indicate, it is unlikely that monitoring, accountability, and effort will have a unidirectional relationship that can be realized across different contexts. Thus, instead of general calls for accountability, it is important that policymakers and managers take stock of the lessons of the implementation literature and think through the range of expected behavioral responses to their policies on the part of street-level bureaucrats, as well as the contextual conditions that will shape those responses.

Note

1. See Zarychta et al. (2020) for a discussion of public service motivation in this same case.

Acknowledgements

We are thankful for the support and assistance we received from staff at the Ministry of Health and the Inter-American Development Bank in Honduras. We are especially grateful for comments and suggestions on the paper from Jane Menken, Elisabeth Root, and Adriana Molina. All errors and omissions are our own.

Disclosure statement

The authors have no conflicts of interest to declare with respect to the research described in this article.

Funding

This work was funded by Division of Graduate Education; SBE Office of Multidisciplinary Activities. This project was completed with financial support from the National Science Foundation (award numbers DGE-1144083 and SMA-1328688) and the Social Science Research Council.

Notes on contributors

Alan Zarychta (azarychta@uchicago.edu) is assistant professor in the Crown Family School of Social Work, Policy, and Practice at the University of Chicago and fellow in the school's Center for Health Administration Studies. His research focuses on comparative public administration and the politics of social services in developing countries, especially in the areas of public health and environmental policy.

Tara Grillos (tgrillos@purdue.edu) is assistant professor in the Department of Political Science at Purdue University and a faculty affiliate of the Purdue Center for the Environment. Her research focuses on human decision-making and behavior in the context of sustainable development, particularly questions of participation, deliberation, collective action, and public goods provision.

Krister P. Andersson (krister.andersson@colorado.edu) is professor in Department of Political Science, faculty associate in the Institute of Behavioral Science, and director of the Center for the Governance of Natural Resources at the University of Colorado Boulder. His research focuses on governmental reforms to address social and environmental problems in developing countries, with a particular interest in local institutional arrangements and the role they play in explaining policy outcomes.

ORCID

Alan Zarychta http://orcid.org/0000-0002-4342-2356
Tara Grillos http://orcid.org/0000-0001-5743-3825
Krister P. Andersson http://orcid.org/0000-0002-9320-8155

References

Abimbola, S., L. Baatiema, and M. Bigdeli. 2019. "The Impacts of Decentralization on Health System Equity, Efficiency and Resilience: A Realist Synthesis of the Evidence." *Health Policy and Planning* 34(8):605–17. doi: 10.1093/heapol/czz055.

Ackerman, J. 2004. "Co-Governance for Accountability: Beyond 'Exit' and 'Voice'." *World Development* 32(3): 447–63. doi: 10.1016/j.worlddev.2003.06.015.

Andersson, K., C. Gibson, and F. Lehoucq. 2006. "Municipal Politics and Forest Governance: Comparative Analysis of Decentralization in Bolivia and Guatemala." *World Development* 34(3):576–95. doi: 10.1016/j.worlddev.2005.08.009.

Andrews, A. C., R. A. Clawson, B. M. Gramig, and L. Raymond. 2013. "Why Do Farmers Adopt Conservation Tillage? An Experimental Investigation of Framing Effects." *Journal of Soil and Water Conservation* 68(6): 501–11. doi: 10.2489/jswc.68.6.501.

Ashraf, N., O. Bandiera, and B. K. Jack. 2014. "No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery." *Journal of Public Economics* 120:1–17. doi: 10.1016/j.jpubeco.2014.06.014.

Banuri, S., and P. Keefer. 2016. "Pro-Social Motivation, Effort and the Call to Public Service." *European Economic Review* 83:139–64. doi: 10.1016/j.euroecorev.2015.10.011.

Battaglio, R. P., P. Belardinelli, N. Bellé, and P. Cantarelli. 2019. "Behavioral Public Administration Ad Fontes: A Synthesis of Research on Bounded Rationality, Cognitive Biases, and Nudging in Public Organizations." *Public Administration Review* 79(3):304–20. doi: 10.1111/puar.12994.

Bauhr, M., and M. Grimes. 2014. "Indignation or Resignation: The Implications of Transparency for Societal Accountability." *Governance* 27(2):291–320. doi: 10.1111/gove.12033.

Banerjee, A., and E. Duflo. 2006. "Addressing Absence." *The Journal of Economic Perspectives: A Journal of the American Economic Association* 20(1):117–32. doi: 10.1257/089533006776526139.

Banerjee, A. V., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani. 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy* 2(1):1–30. doi: 10.1257/pol.2.1.1.

Bellé, N. 2013. "Experimental Evidence on the Relationship between Public Service Motivation and Job Performance." *Public Administration Review* 73(1):143–53. doi: 10.1111/j.1540-6210.2012.02621.x.

Björkman, M., and J. Svensson. 2010. "When is Community-Based Monitoring Effective? Evidence from a Randomized Experiment in Primary Health in Uganda." *Journal of the European Economic Association* 8(2–3): 571–81. doi: 10.1111/j.1542-4774.2010.tb00527.x.

Bovens, M. 2007. "Public Accountability." In *The Oxford Handbook of Public Management*, edited by E. Ferlie, L. E. Lynn Jr., and C. Pollitt, 182–208. New York: Oxford University Press.

Brehm, J., and S. Gates. 2015. "Bureaucratic Politics Arising from, Not Defined by, a Principal–Agency Dyad." *Journal of Public Administration Research and Theory* 25(1):27–42. doi: 10.1093/jopart/muu045.

Brinkerhoff, D. W. 2004. "Accountability and Health Systems: Toward Conceptual Clarity and Policy Relevance." *Health Policy and Planning* 19(6):371–9. doi: 10.1093/heapol/czh052.

Brinkerhoff, D. W., and T. J. Bossert. 2014. "Health Governance: Principal-Agent Linkages and Health System Strengthening." *Health Policy and Planning* 29(6):685–93. doi: 10.1093/heapol/czs132.

Brinkerhoff, D. W., and A. Wetterberg. 2016. "Gauging the Effects of Social Accountability on Services, Governance, and Citizen Empowerment." *Public Administration Review* 76(2):274–86. doi: 10.1111/puar.12399.

Brodkin, E. Z. 2007. "Bureaucracy Redux: Management Reformism and the Welfare State." *Journal of Public Administration Research and Theory* 17(1):1–17. doi: 10.1093/jopart/muj019.

Brodkin, E. Z. 2008. "Accountability in Street-Level Organizations." *International Journal of Public Administration* 31(3):317–36. doi: 10.1080/01900690701590587.

Brodkin, E. Z. 2011. "Policy Work: Street-Level Organizations under New Managerialism." *Journal of Public Administration Research and Theory* 21(Suppl. 2):i253–77. doi: 10.1093/jopart/muq093.

Cabral, S., and M.-F. Santos. 2018. "Accountability Mechanisms in Public Services: Activating New Dynamics in a Prison System." *International Public Management Journal* 21(5):795–821. doi: 10.1080/10967494.2016.1141815.

Cammett, M., and L. MacLean. 2011. "Introduction: The Political Consequences of Non-State Social Welfare in the Global South." *Studies in Comparative International Development* 46(1):1–21. doi: 10.1007/s12116-010-9083-7.

Cárdenas, J.-C., and E. Ostrom. 2004. "What Do People Bring into the Game? Experiments in the Field about Cooperation in the Commons." *Agricultural Systems* 82(3):307–26. doi: 10.1016/j.agsy.2004.07.008.

Carpenter, J., and E. Gong. 2016. "Motivating Agents: How Much Does the Mission Matter?" *Journal of Labor Economics* 34(1):211–36. doi: 10.1086/682345.

Carrigan, C., S. K. Pandey, and G. G. Van Ryzin. 2020. "Pursuing Consilience: Using Behavioral Public Administration to Connect Research on Bureaucratic Red Tape, Administrative Burden, and Regulation." *Public Administration Review* 80(1):46–52. doi: 10.1111/puar.13143.

Carsey, T. M., and J. J. Harden. 2013. *Monte Carlo Simulation and Resampling Methods for Social Science*. Los Angeles, CA: SAGE Publications, Inc.
Channa, A., and J.-P. Faguet. 2016. "Decentralization of Health and Education in Developing Countries: A Quality-Adjusted Review of the Empirical Literature." *The World Bank Research Observer* 31(2):199–241. doi: 10.1093/wbro/lkw001.
Chong, A., A. L. De La O, D. Karlan, and L. Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification." *The Journal of Politics* 77(1):55–71. doi: 10.1086/678766.
Chong, D., and J. N. Druckman. 2007. "Framing Theory." *Annual Review of Political Science* 10(1):103–26. doi: 10.1146/annurev.polisci.10.072805.103054.
Christensen, R. K., L. Paarlberg, and J. L. Perry. 2017. "Public Service Motivation Research: Lessons for Practice." *Public Administration Review* 77(4):529–42. doi: 10.1111/puar.12796.
Cuadrado-Ballesteros, B. 2014. "The Impact of Functional Decentralization and Externalization on Local Government Transparency." *Government Information Quarterly* 31(2):265–77. doi: 10.1016/j.giq.2013.10.012.
Cucciniello, M., G. A. Porumbescu, and S. Grimmelikhuijsen. 2017. "25 Years of Transparency Research: Evidence and Future Directions." *Public Administration Review* 77(1):32–44. doi: 10.1111/puar.12685.
Dasandi, N., and M. Esteve. 2017. "The Politics–Bureaucracy Interface in Developing Countries." *Public Administration and Development* 37(4):231–45. doi: 10.1002/pad.1793.
de Boer, N., and J. Eshuis. 2018. "A Street-Level Perspective on Government Transparency and Regulatory Performance: Does Relational Distance Matter?" *Public Administration* 96(3):452–67. doi: 10.1111/padm.12517.
Deci, E. L., and R. M. Ryan. 2000. "The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior." *Psychological Inquiry* 11(4):227–68. doi: 10.1207/S15327965PLI1104_01.
DellaVigna, S., J. A. List, U. Malmendier, and G. Rao. 2016. *Estimating Social Preferences and Gift Exchange at Work* (Working Paper). National Bureau of Economic Research.
Dhaliwal, I., and R. Hanna. 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics* 124:1–21. doi: 10.1016/j.jdeveco.2016.08.008.
Dubnick, M. 2005. "Accountability and the Promise of Performance: In Search of the Mechanisms." *Public Performance & Management Review* 28(3):376–417.
Escobar-Lemmon, M., and A. D. Ross. 2014. "Does Decentralization Improve Perceptions of Accountability? Attitudinal Evidence from Colombia." *American Journal of Political Science* 58(1):175–88. doi: 10.1111/ajps.12043.
Etzioni, A. 2010. "Is Transparency the Best Disinfectant?" *Journal of Political Philosophy* 18(4):389–404. doi: 10.1111/j.1467-9760.2010.00366.x.
Etzioni, A. 2014. "The Limits of Transparency." *Public Administration Review* 74(6):687–8. doi: 10.1111/puar.12276.
Ewert, B. 2020. "Focusing on Quality Care Rather than 'Checking Boxes': How to Exit the Labyrinth of Multiple Accountabilities in Hybrid Healthcare Arrangements." *Public Administration* 98(2):308–24. doi: 10.1111/padm.12556.
Falk, A., and A. Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics* 24(1):39–57. doi: 10.1086/497818.
Ferry, L., and P. Eckersley. 2015. "Accountability and Transparency: A Nuanced Response to Etzioni." *Public Administration Review* 75(1):11–2. doi: 10.1111/puar.12303.
Finan, F., B. A. Olken, and R. Pande. 2015. *The Personnel Economics of the State* (Working Paper). National Bureau of Economic Research.
Franco, L. M., S. Bennett, and R. Kanfer. 2002. "Health Sector Reform and Public Sector Health Worker Motivation: A Conceptual Framework." *Social Science & Medicine* (1982) 54(8):1255–66. doi: 10.1016/S0277-9536(01)00094-6.
Gailmard, S. 2009. "Multiple Principals and Oversight of Bureaucratic Policy-Making." *Journal of Theoretical Politics* 21(2):161–86. doi: 10.1177/0951629808100762.
Gailmard, S. 2010. "Politics, Principal–Agent Problems, and Public Service Motivation." *International Public Management Journal* 13(1):35–45. doi: 10.1080/10967490903547225.
Gailmard, S., and J. W. Patty. 2012. "Formal Models of Bureaucracy." *Annual Review of Political Science* 15(1):353–77. doi: 10.1146/annurev-polisci-031710-103314.
Gill, D., and V. L. Prowse. 2011. *A Novel Computerized Real Effort Task Based on Sliders* (SSRN Scholarly Paper). Rochester, NY: Social Science Research Network. Retrieved August 31, 2018 (https://papers.ssrn.com/abstract=1732324).
Grillos, T., A. Zarychta, and K. Andersson. 2021. "Governance Reform, Decentralization, and Teamwork in Public Service Delivery: Evidence from the Honduran Health Sector." *Public Administration* :1–27. doi: 10.1111/padm.12722.

Grindle, M. S. 2004. "Good Enough Governance: Poverty Reduction and Reform in Developing Countries." *Governance* 17(4):525–48. doi: 10.1111/j.0952-1895.2004.00256.x.
Grossman, G. 2011. "Lab-in-the-Field Experiments." *Newsletter of the APSA Experimental Section* 2(2):13–9.
Han, Y., and J. L. Perry. 2020. "Employee Accountability: Development of a Multidimensional Scale." *International Public Management Journal* 23(2):224–51. doi: 10.1080/10967494.2019.1690606.
Harrison, G. W., and J. A. List. 2004. "Field Experiments." *Journal of Economic Literature* 42(4):1009–55. doi: 10.1257/0022051043004577.
Heinrich, C. J., and G. Marschke. 2010. "Incentives and Their Dynamics in Public Sector Performance Management Systems." *Journal of Policy Analysis and Management* 29(1):183–208. doi: 10.1002/pam.20484.
Hill, M., and P. Hupe. 2006. "Analysing Policy Processes as Multiple Governance: Accountability in Social Policy." *Policy & Politics* 34(3):557–73. doi: 10.1332/030557306777695280.
Hood, C. 1991. "A Public Management for All Seasons?" *Public Administration* 69(1):3–19. doi: 10.1111/j.1467-9299.1991.tb00779.x.
Hupe, P., and M. Hill. 2007. "Street-Level Bureaucracy and Public Accountability." *Public Administration* 85(2):279–99. doi: 10.1111/j.1467-9299.2007.00650.x.
Imbens, G. W., and D. B. Rubin. 2015. *"Causal Inference for Statistics." Social, and Biomedical Sciences*. New York, NY: Cambridge University Press.
James, O., S. R. Jilke, and G. G. Van Ryzin. 2017. "Behavioural and Experimental Public Administration: Emerging Contributions and New Directions." *Public Administration* 95(4):865–73. doi: 10.1111/padm.12363.
Jayal, N. G. 2007. "New Directions in Theorising Social Accountability?" *IDS Bulletin* 38(6):105–12. doi: 10.1111/j.1759-5436.2007.tb00425.x.
Joaquin, M. E., and T. J. Greitens. 2011. "The Accountability-Performance Link." *Public Performance & Management Review* 34(3):323–49. doi: 10.2753/PMR1530-9576340301.
Konow, J. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." American Economic Review 90(4):1072–91. doi: 10.1257/aer.90.4.1072.
Lancaster, T. D. 2014. "Political Monitoring' as an Analytical Concept: A Lasting Legacy?" *Publius: The Journal of Federalism* 44(2):249–74. doi: 10.1093/publius/pju005.
Lerner, J. S., and P. E. Tetlock. 1999. "Accounting for the Effects of Accountability." *Psychological Bulletin* 125(2):255–75. doi: 10.1037/0033-2909.125.2.255.
Lieberman, E. S. 2011. "The Perils of Polycentric Governance of Infectious Disease in South Africa." *Social Science & Medicine (1982)* 73(5):676–84. doi: 10.1016/j.socscimed.2011.06.012.
Lipsky, M. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. New York: Russell Sage Foundation.
Manning, N. 2001. "The Legacy of the New Public Management in Developing Countries." *International Review of Administrative Sciences* 67(2):297–312. doi: 10.1177/0020852301672009.
May, P. J., and S. C. Winter. 2009. "Politicians, Managers, and Street-Level Bureaucrats: Influences on Policy Implementation." *Journal of Public Administration Research and Theory* 19(3):453–76. doi: 10.1093/jopart/mum030.
Meier, K. J., and L. J. O'Toole. 2006. "Political Control versus Bureaucratic Values: Reframing the Debate." *Public Administration Review* 66(2):177–92. doi: 10.1111/j.1540-6210.2006.00571.x.
Meijer, A. 2013. "Understanding the Complex Dynamics of Transparency." *Public Administration Review* 73(3):429–39. doi: 10.1111/puar.12032.
Mokdad, A. H., K. E. Colson, P. Zúñiga-Brenes, D. Ríos-Zertuche, E. B. Palmisano, E. Alfaro-Porras, B. W. Anderson, M. Borgo, S. Desai, M. C. Gagnier, et al. 2015. "Salud Mesoamérica 2015 Initiative: Design, Implementation, and Baseline Findings." *Population Health Metrics* 13(1):3. doi: 10.1186/s12963-015-0034-4.
Morton, R. B., and K. C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.
Moynihan, D. P. 2010. "A Workforce of Cynics? The Effects of Contemporary Reforms on Public Service Motivation." *International Public Management Journal* 13(1):24–34. doi: 10.1080/10967490903547167.
Moynihan, D. P., and S. K. Pandey. 2007. "The Role of Organizations in Fostering Public Service Motivation." *Public Administration Review* 67(1):40–53. doi: 10.1111/j.1540-6210.2006.00695.x.
Mulgan, R. 2000. "Accountability': An Ever-Expanding Concept?" *Public Administration* 78(3):555–73. doi: 10.1111/1467-9299.00218.
Nelson-Nuñez, J. 2019. "Substitution or Facilitation: Service-Delivery NGOs and Political Engagement in the Peruvian Amazon." *Comparative Political Studies* 52(3):445–77. doi: 10.1177/0010414018774376.
Oliveros, V., and C. Schuster. 2018. "Merit, Tenure, and Bureaucratic Behavior: Evidence from a Conjoint Experiment in the Dominican Republic." *Comparative Political Studies* 51(6):759–92. doi: 10.1177/0010414017710268.
Olken, B. A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115(2):200–49. doi: 10.1086/517935.

Ospina, S., N. C. Grau, and A. Zaltsman. 2004. "Performance Evaluation, Public Management Improvement and Democratic Accountability." *Public Management Review* 6(2):229–51. doi: 10.1080/1471903042000189119.
Ossege, C. 2012. "Accountability – Are We Better off without It?" *Public Management Review* 14(5):585–607. doi: 10.1080/14719037.2011.642567.
Ostrom, E. 2000. "Crowding out Citizenship." *Scandinavian Political Studies* 23(1):3–16.
Pandey, S. K., and E. C. Stazyk. 2008. "Antecedents and Correlates of Public Service Motivation." Pp. 101–17, in *Motivation in Public Management: The Call of Public Service*, edited by J. L. Perry and A. Hondeghem. New York: Oxford University Press Inc.
Pedersen, M. J. 2015. "Activating the Forces of Public Service Motivation: Evidence from a Low-Intensity Randomized Survey Experiment." *Public Administration Review* 75(5):734–46. doi: 10.1111/puar.12325.
Perry, J. L. 1996. "Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity." *Journal of Public Administration Research and Theory* 6(1):5–22. doi: 10.1093/oxfordjournals.jpart.a024303.
Perry, J. L., and L. R. Wise. 1990. "The Motivational Bases of Public Service." *Public Administration Review* 50(3): 367–73. doi: 10.2307/976618.
Piatak, J., Z. Mohr, and S. Leland. 2017. "Bureaucratic Accountability in Third-Party Governance: Experimental Evidence of Blame Attribution during Times of Budgetary Crisis." *Public Administration* 95(4):976–89. doi: 10.1111/padm.12341.
Ramos, C., and A. Milanesi. 2020. "The Neo-Weberian State and the Neodevelopmentalist Strategies in Latin America: The Case of Uruguay." *International Review of Administrative Sciences* 86(2):261–77. doi: 10.1177/0020852318763525.
Rasinski, K. A. 1989. "The Effect of Question Wording on Public Support for Government Spending." *Public Opinion Quarterly* 53(3):388–94. doi: 10.1086/269158.
Rasul, I., and D. Rogger. 2018. "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service." *The Economic Journal* 128(608):413–46. doi: 10.1111/ecoj.12418.
Resh, W. G., J. D. Marvel, and B. Wen. 2018. "The Persistence of Prosocial Work Effort as a Function of Mission Match." *Public Administration Review* 78(1):116–25. doi: 10.1111/puar.12882.
Root, E. D., A. Zarychta, B. Bermudez Tapia, T. Grillos, K. Andersson, and J. Menken. 2020. "Organizations Matter in Local Governance: Evidence from Health Sector Decentralization in Honduras." *Health Policy and Planning* 35 (9):1168–79. doi: 10.1093/heapol/czaa084.
Romzek, B. S. 2000. "Dynamics of Public Sector Accountability in an Era of Reform." *International Review of Administrative Sciences* 66(1):21–44. doi: 10.1177/0020852300661004.
Romzek, B. S., and M. J. Dubnick. 1987. "Accountability in the Public Sector: Lessons from the Challenger Tragedy." *Public Administration Review* 47(3):227–38. doi: 10.2307/975901.
Ruhl, J. M. 2010. "Honduras Unravels." *Journal of Democracy* 21:93–107.
Schillemans, T., M. Van Twist, and I. Vanhommerig. 2013. "Innovations in Accountability." *Public Performance & Management Review* 36(3):407–35. doi: 10.2753/PMR1530-9576360302.
Serra, D. 2012. "Combining Top-Down and Bottom-Up Accountability: Evidence from a Bribery Experiment." Journal of Law, Economics, and Organization 28(3):569–87. doi: 10.1093/jleo/ewr010.
Siddiqi, S., T. I. Masud, S. Nishtar, D. H. Peters, B. Sabri, K. M. Bile, and M. A. Jama. 2009. "Framework for Assessing Governance of the Health System in Developing Countries: Gateway to Good Governance." *Health Policy (Amsterdam, Netherlands)* 90(1):13–25. doi: 10.1016/j.healthpol.2008.08.005.
Smith, J. 2016. "The Motivational Effects of Mission Matching: A Lab-Experimental Test of a Moderated Mediation Model." *Public Administration Review* 76(4):626–37. doi: 10.1111/puar.12514.
Sniderman, P. M., and S. M. Theriault. 2004. "The Structure of Political Argument and the Logic of Issue Framing." Pp. 133–65, in *Studies in Public Opinion*, edited by. W. E. Saris and P. M. Sniderman. Princeton, NJ: Princeton University Press.
Treisman, D. 2007. *The Architecture of Government: Rethinking Political Decentralization*. New York, NY: Cambridge University Press.
Tversky, A., and D. Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science (New York, N.Y.)* 211(4481):453–8. doi: 10.1126/science.7455683.
Vandenabeele, W., and S. Van de Walle. 2008. "International Difference in Public Service Motivation: Comparing Regions across the World." Pp. 223–44, in *Motivation in Public Management: The Call of Public Service*, edited by J. L. Perry and A. Hondeghem. New York: Oxford University Press.
Véron, R., G. Williams, S. Corbridge, and M. Srivastava. 2006. "Decentralized Corruption or Corrupt Decentralization? Community Monitoring of Poverty-Alleviation Schemes in Eastern India." *World Development* 34(11):1922–41. doi: 10.1016/j.worlddev.2005.11.024.
Wright, B. E. 2004. "The Role of Work Context in Work Motivation: A Public Sector Application of Goal and Social Cognitive Theories." *Journal of Public Administration Research and Theory* 14(1):59–78. doi: 10.1093/jopart/muh004.

Wright, B. E. 2007. "Public Service and Motivation: Does Mission Matter?" *Public Administration Review* 67(1): 54–64. doi: 10.1111/j.1540-6210.2006.00696.x.

Wright, B. E., and S. K. Pandey. 2011. "Public Organizations and Mission Valence: When Does Mission Matter?" *Administration & Society* 43(1):22–44. doi: 10.1177/0095399710386303.

Wright, G. D., K. P. Andersson, C. C. Gibson, and T. P. Evans. 2016. "Decentralization Can Help Reduce Deforestation When User Groups Engage with Local Government." *Proceedings of the National Academy of Sciences of the United States of America* 113(52):14958–63. doi: 10.1073/pnas.1610650114.

Yanez-Pagans, M., and G. Machicado-Salas. 2014. "Bureaucratic Delay, Local-Level Monitoring, and Delivery of Small Infrastructure Projects: Evidence from a Field Experiment in Bolivia." *World Development* 59:394–407. doi: 10.1016/j.worlddev.2014.02.004.

Zarychta, A. 2020. "Making Social Services Work Better for the Poor: Evidence from a Natural Experiment with Health Sector Decentralization in Honduras." *World Development* 133:104996. doi: 10.1016/j.worlddev.2020.104996.

Zarychta, A., K. P. Andersson, E. D. Root, J. Menken, and T. Grillos. 2019a. "Assessing the Impacts of Governance Reforms on Health Services Delivery: A Quasi-Experimental, Multi-Method, and Participatory Approach." *Health Services and Outcomes Research Methodology* 19(4):241–58. doi: 10.1007/s10742-019-00201-8.

Zarychta, A., K. P. Andersson, E. D. Root, J. Menken, and T. Grillos. 2019b. "Supplemental Appendix for 'Assessing the Impacts of Governance Reforms on Health Services Delivery: A Quasi-Experimental, Multi-Method, and Participatory Approach'." *Health Services and Outcomes Research Methodology* 19(4):241–58. doi: 10.1007/s10742-019-00201-8.

Zarychta, A., T. Grillos, and K. P. Andersson. 2020. "Public Sector Governance Reform and the Motivation of Street-Level Bureaucrats in Developing Countries." *Public Administration Review* 80(1):75–91. doi: 10.1111/puar.13132.