

Governance reform, decentralization, and teamwork in public service delivery: Evidence from the Honduran health sector

Tara Grillos¹  | Alan Zarychta²  | Krister Andersson³ 

¹Political Science, Purdue University, West Lafayette, Indiana, USA

²School of Social Service Administration, University of Chicago, Chicago, Illinois, USA

³Institute of Behavioral Science, University of Colorado Boulder, Boulder, Colorado, USA

Correspondence

Tara Grillos, Political Science, Purdue University, West Lafayette, IN.
Email: tgrillos@purdue.edu

Abstract

Public service delivery improves when civil servants work together effectively as teams. While decentralization reforms are common strategies for enhancing the delivery of health services in developing countries, most studies emphasize their effects through rational-choice mechanisms. Fewer studies consider the behavioral implications of decentralization and its potential to improve or hinder the day-to-day work environment for health sector staff. We use an incentivized behavioral game to assess the effectiveness of teamwork among civil servants in decentralized and comparable centrally administered municipal health systems in Honduras. We find that teams from decentralized municipalities are less effective at working together and that this is driven by the adverse effects of the reform among existing staff: new staff pairs are relatively effective, while existing staff do not work as well with each other and in mixed teams. Our findings suggest that policymakers should take measures to help ensure that governance reforms do not undermine the effectiveness of existing staff.

Governments often engage in institutional reforms with the goal of improving public sector performance. These reforms commonly try to reshape governance, namely, the public decision-making processes through which socially binding agreements are created and enforced (Root et al., 2020; Andersson et al., 2009). For example, decentralization, which is one of the most widely implemented governance reforms of the last several decades, shifts power and responsibilities, and sometimes resources as well, away from central organizations and toward more local organizations (Faguet, 2014; Rondinelli et al., 1989). This is often done in an effort to improve accountability and thereby produce better public services for local communities (Blair, 2000; Ribot, 2002; Faguet, 2004, 2012; Diamond &

Tsalik, 1999; Stepan, 2001). Incentive-based governance reforms, such as performance-based financing and results-based aid, have also emerged as popular policies to encourage improved administrative capacity in the public sector (Brinkerhoff & Brinkerhoff, 2015; El Bcheraoui et al., 2017; Hood, 1991; Loevinsohn & Harding, 2005; Manning, 2001). While accountability and incentive mechanisms are clearly important, a large body of social science research demonstrates that programs intended to operate through these rational-choice mechanisms often have additional social and psychological effects on target populations (Bowles, 2008; Deci et al., 1999; Frey, 1994; Gneezy & Rustichini, 2000; Heyman & Ariely, 2004).

Ultimately, street-level bureaucrats are the targets of these governance reforms as they reside at the end of the service delivery chain (May & Winter, 2009). Frontline service providers and local public officials play a crucial role in the implementation of government policy and public services delivery (Brodkin, 1990; Lipsky, 1980), and they exercise a great degree of influence over how they carry out their daily activities (Brodkin, 2008; Hupe & Hill, 2007; Le Grand, 2010; May & Winter, 2009; Scholz et al., 1991). In the developing country context, where weak state capacity and limited resources constrain the extent of direct accountability, bureaucrats can have even greater discretion (Jiang, 2018) and play a more important role in directly shaping local outcomes (Andersson, 2004; Ang, 2017; Geddes, 1994; Larbi, 2001; Tankha, 2009).

Governance reforms affect not only material incentives faced by public service providers but also the choice architecture within which they make decisions (Brodkin, 2008; Brodkin, 2011; Moynihan, 2010; Moynihan & Soss, 2014). The growing literature on behavioral public administration emphasizes micro-level foundations of individual behavior (Battaglio Jr et al., 2019; Grimmelikhuijsen et al., 2016; Moynihan, 2010). However, much of this work has focused on small-scale 'nudges,' rather than the behavioral impact of broad-scale policy interventions. Recent reviews of this literature urge us to move beyond nudges toward "behavioral science research that focuses on meso- and macro-level understandings of policies and institutions" (Bhanot & Linos, 2020, p. 169) and to engage more with outcomes of interest to public management (Hassan & Wright, 2019).

We examine the impact of meso-level governance reform on teamwork, namely, the ability of public service providers to coordinate their activities effectively during the implementation of public policies and delivery of social services. We use an incentivized behavioral game to measure and compare successful teamwork in both decentralized and centrally administered municipal health systems in Honduras. We find that decentralization reduces the effectiveness of teamwork and attribute this negative net effect to the how existing staff members, hired before the reform, have experienced and reacted to the reform.

1 | THE IMPORTANCE OF TEAMWORK FOR HEALTH SERVICES

Social service delivery is enhanced when bureaucrats, civil servants, and frontline service providers are able to work together effectively and coordinate their joint activities. For example, teamwork is critical to develop an annual operating plan, design and carry out targeted interventions, or come together as a team to solve a difficult problem faced by a client. The study of teamwork is critically important for effective service delivery but has been underemphasized in work on governance reform (Baker et al., 2006; Rigoli & Dussault, 2003). Although there is a large and diverse literature on collaboration (Ansell & Gash, 2018; Emerson et al., 2012), that scholarship focuses on interactions among organizations, while teamwork, as defined below in the context of public administration, is about micro-level interactions between individual public servants within and across levels of government, working to implement policy directives.

Teamwork refers to "groups of two or more individuals working interdependently toward a shared goal" (Salas et al., 1992; Weaver et al., 2012). Multidisciplinary teamwork is especially important in primary health care, which is characterized by complex problems that require different types of expertise (Solheim et al., 2007). A "complex task" is one in which there is no single path to the solution (Campbell, 1988), and so, coordinating strategies across

multiple actors is essential. Complex tasks are also more difficult to monitor and incentivize (Dixit, 2002; Jakobsen et al., 2018), making behavioral aspects of performance that much more salient.

In the health sector, teams of doctors, nurses, social workers, and administrators routinely interface when diagnosing, treating, and following up on individual patient cases. When developing outreach and promotion campaigns to address public health concerns, similar teamwork is required. Such coordination is essential across levels within the health sector hierarchy as well. For example, local health workers may know which supplies are in the highest demand at their own health centers, while regional administrators will know more about the costs and seasonal supply availabilities of different medicines and materials. Both pieces of information are relevant to making efficient decisions about medical supply provision, and good communication and coordination across levels are thus required to be successful in delivering the required public health services.

Simply working in groups does not necessarily translate into successful teamwork (Finn et al., 2010; Rigoli & Dussault, 2003). Prior work identifies communication and coordination as two of the critical components of successful teamwork (Weaver et al., 2012). Communication is the exchange of information that enables groups to perform tasks (Fussel et al., 1998), and coordination is a “process of orchestrating the sequence and timing of interdependent actions” (Marks et al., 2001, p. 367). Research on healthcare settings in the United States suggests communication and coordination are often lacking, contributing to lower-quality services (Baldwin & Daugherty, 2008; Danjoux Meth et al., 2009; Johnson, 2009). Prior research has also emphasized the potential role of leadership (Baker et al., 2006; Leggat, 2007).

Existing work has identified three major approaches to facilitating effective teamwork: select team members who already coordinate well, modify contextual conditions to support better teamwork, or build teamwork competencies through training (Baker et al., 2006, p. 1584). Training has been the most commonly used strategy (Baker et al., 2006), but governance reforms, and the hybrid service delivery models they produce, have high potential to influence both selection and context.

2 | GOVERNANCE REFORM IN LATIN AMERICA

Public service delivery in developing countries is often lacking, with large variation across administrative subunits (Andersson et al., 2009). In response, governments have frequently engaged in institutional reform efforts with the goal of improving performance (World Bank, 2004). Two of the most common reform types used in practice are decentralization (Faguet, 2014) and performance-based management (Moynihan & Ingraham, 2003).

Beginning in the 1980s, the New Public Management (NPM) movement motivated a wave of reforms modeled after the private sector (Batley, 1999; Hood, 1991). NPM refers to a loose collection of distinct ideas (Manning, 2001), but NPM philosophy was closely linked to promarket discourse and, often, privatization (Pollitt et al., 2007). Results of such reforms have varied substantially across settings, with the most successful efforts involving goal integration, political support, and leadership (Moynihan & Ingraham, 2003). Two key features of NPM as implemented in Latin America were increased autonomy for public administrators and accountability based on performance rather than process (Arellano-Gault & Ramírez, 2008).

NPM was closely linked to decentralization in this region, with the two reform types seen as complementary in theory (Arellano-Gault, 2000) and often linked in practice (Hope Sr & Chikulo, 2000). Decentralization aims to shift decision-making authority to lower levels within the government structure, closer to the site of implementation (Faguet, 2004; Rondinelli et al., 1989). Proponents argue that these more proximate decision makers can make more appropriate choices because they have better information about local conditions and are more easily held accountable by local actors (Diamond & Tsalik, 1999; Hayek, 1945; Oates, 1977).

There are mixed results regarding the effect of decentralization on health outcomes (Bossert et al., 2003; Campos & Hellman, 2005; Jeppsson & Okunzi, 2000; Root et al., 2020) but with relatively consistent positive effects on

reducing infant mortality rates (Guanais & Macinko, 2009; Robalino et al., 2001). However, in Latin America, some evidence calls into question the very logic of decentralization. Smaller decision-making units do not necessarily perform better than larger ones (Avellaneda & Gomes, 2015), and too much administrative autonomy can actually undermine equity, leading to undesirable outcomes (Cárdenas & Ramírez de la Cruz, 2017; Galiani et al., 2008). The effects of horizontal, collaborative arrangements vary substantially with the type of collaboration, for example, with municipal associations having opposite effects from formal municipal agreements (Piña & Avellaneda, 2018).

In Latin America, the rise of leftist governments led to a different conceptualization of the public sector that sought to return to the centrality of the state (Boschi & Gaitán, 2009; Gaitán, 2014), in direct contrast to the market-oriented focus that NPM had previously encouraged (Ramos & Milanese, 2020). A second generation of public sector reform in Latin America aimed to professionalize the bureaucracy (Panizza & Philip, 2005). Ramos and Milanese (2020) liken these efforts, as applied within public administration, to the “neo-Weberianism” adopted by many European countries. They argue that, while this neo-Weberian approach has gained popularity in Latin America, institutional deficiencies and political realities in the region led to a public management model that is more accurately described as “imperfect neo-Weberianism” (Ramos & Milanese, 2020). Laws intended to increase the professionalization of the bureaucracy have failed to reduce corruption throughout Latin America (Langbein & Sanabria, 2017), perhaps due to these obstacles.

The Honduran health sector reform examined in this study must be considered in light of this regional history. While principally framed as decentralization reform, it also includes components of performance-based management. Specifically, the new meso-level intermediary managing organizations enter into performance-based agreements with the Ministry of Health, which provide them considerable control over hiring practices at the local level and supervision of those primary care health workers, among other responsibilities. This reform thus has aspects of imperfect neo-Weberianism, alongside some philosophical legacies of NPM.

Considerable research on street-level bureaucracy and implementation, however, shows that, in many cases, performance may be only minimally influenced by efforts to exert supervisory control (e.g., Brehm & Gates, 1999; Brodtkin, 2008; Lipsky, 1980). Managerial strategies can successfully alter organizational norms (Miller, 1992; Riccucci et al., 2004), even in a context characterized by initial resistance to change (Ateş, 2004). Nonetheless, these types of change may also have adverse consequences. Critiques of performance-based management regimes argue that external monitoring and rewards can displace workers' original job motivation and satisfaction (Jacobsen et al., 2014; Jakobsen & Mortensen, 2016; Ryan & Weinstein, 2009; Soss et al., 2011) or undermine social relationships (Magrath & Nichter, 2012; Powell-Jackson et al., 2009), potentially leading to decreased team effectiveness and work performance.

Such health sector reforms also regularly give authority to local actors and organizations to hire, manage, and potentially fire civil servants (Rigoli & Dussault, 2003), thus paving the way for the selection of new employees who lack attachment to the prior way of doing things. At the same time, passing reforms in the first place usually involves a negotiation with the unions representing civil servants. The need to garner political support for reforms can lead to the adoption of policies ‘without losers’ (Panizza, 2004), producing a blend of different public management philosophies coexisting within the same reform. According to Panizza (2004), this can result in a political culture in which resistance to change is inescapable.

Given their considerable discretion, street-level practitioners are positioned with a remarkable ability to resist new ways of doing things (Lipsky, 1980; Maynard-Moody et al., 2003; Meyers et al., 1998; Pressman & Wildavsky, 1979). Long-standing employees may be particularly resistant to change if they are too entrenched in prior practices (Lipsky, 1980, pp. 143–144). In the wake of contemporary governance reforms, a service delivery system is likely to end up with a mix of permanent/protected civil servants and new contracted/temporary civil servants who need to work together to deliver social services to their populations. Thus, effects arising from the selection of new staff coexist with effects on existing staff through changes to contextual features of their work environments. To the extent that health systems represent social systems (Gilson, 2003; Nichter, 1986), reforms require an adjustment period: discarding old

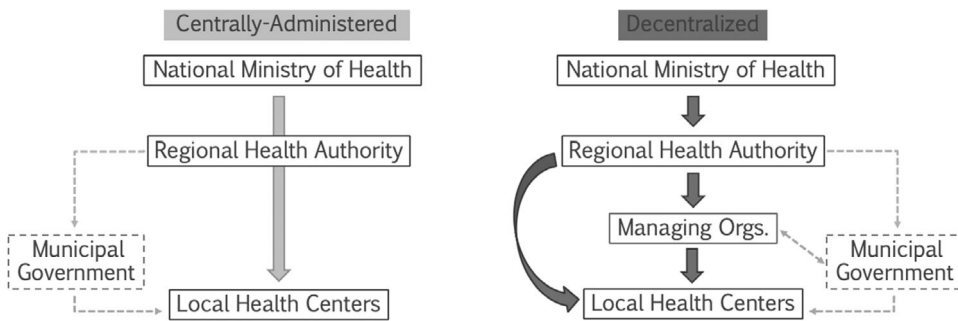


FIGURE 1 Delivery of primary care health Services in Honduras. An earlier version of this figure first appeared in Zarychta et al. (2019)

work patterns, instituting/negotiating new ones, and socializing and incorporating new individuals. During this adjustment period, resistance to change among preexisting staff is likely to lead to discouragement and perhaps conflict, with potential effects on teamwork.

3 | THE SETTING: THE HONDURAN HEALTH SECTOR

Beginning in 2007, Honduras began to roll out a governance reform within the health sector. Like many others, the Honduran reform involves elements of both decentralization and performance-based management, as illustrated by its two key components (Figure 1). First, it deconcentrates responsibility for oversight from the national Ministry of Health to departmental-level regional health authorities. Second, it devolves responsibility for day-to-day service delivery to local managing organizations through performance-based agreements (MOH, 2009, 2010; Zarychta, 2020). Local managing organizations may be nongovernmental organizations, individual municipal governments, or associations of municipalities. Public health centers at the local level act as the direct health services provider for the majority of Hondurans. Workers at these centers report to the regional health authority in the traditional, centrally administered setting, but those workers in the decentralized setting report instead to the new local managing organization.

This reform changes the contextual conditions within which public servants work and also introduces incentives for selection of high-quality, new staff. While hiring decisions were the responsibility of the central Ministry of Health through its Regional Health Authorities prior to the decentralization reform, the reform transferred a large portion of that responsibility to local managing organizations. Furthermore, the contracts between these managing organizations and the Ministry of Health are designed from a “results-based management” perspective. They define a set of performance targets related to the production of key health services (with an emphasis on child and maternal health), and the managing organizations are assessed (and potentially penalized) based on whether or not they are making adequate progress toward meeting those goals (Carmenate-Milián et al., 2017; Garcia Prado & Peña, 2010; MOH, 2009, 2018). Thus, it is in the interest of the managing organizations to identify and hire staff that they believe will effectively contribute to the delivery and uptake of key health services.

The Honduran reform was rolled out in a staggered manner. Some municipalities had already received the reform, while others remained under the centralized form of administration at the time of our study. This made the Honduran context an excellent setting for examining the social and behavioral impacts of this commonly used suite of reforms.

4 | RESEARCH METHODS

4.1 | Quasiexperimental approach

The Honduran Ministry of Health has a nationwide plan for implementing this reform, but at the time of our study, it had been implemented in only 90 of 298 total municipalities. Selection of municipalities for the initial waves of reform was not random and was instead based partially on need (Zarychta, 2020). To address nonrandom selection, we used a common propensity score approach to select a matched sample for data collection (Imbens & Rubin, 2015). The propensity score was developed using 35 potential covariates related to treatment (e.g., implementation of the decentralization reform at the municipal level) reflecting socioeconomic conditions, health system performance, demographic characteristics, and political factors. This type of quasiexperimental design provides us with plausible treatment and control groups for analytical comparisons (Dunning, 2012).

Taking advantage of this, we identified a subset of 31 comparable rural municipalities, 21 of which have already received the decentralization reform and 10 of which are still operating under the centrally administered structure (see Figure 2). Our study focused on rural areas specifically because they are traditionally underserved, and thus, improvements were one main goal of the reform. At the time of data collection, all treatment municipalities had been under the reform for a minimum of 6 years. We ensured that the sample would include representation of distinct managing organization types and that no more than three municipalities would be selected from a single department because of the limited number of department-level officials occupying the roles we sought to recruit into our workshops. Subject to those constraints, we selected the subset of municipalities that matched best based on the propensity score. We also include weights based on that propensity score in all regression analyses associated with this research.

Full details of the matching analysis are published in Zarychta et al. (2019). Among the decentralized municipalities that we selected for inclusion in the study workshop, the average propensity score predicting decentralization was 0.64. Within those we selected for our control group, the average score was 0.58 (an improvement from the prematching control group score of 0.44). While balance is reasonable between the groups, there are a few high-

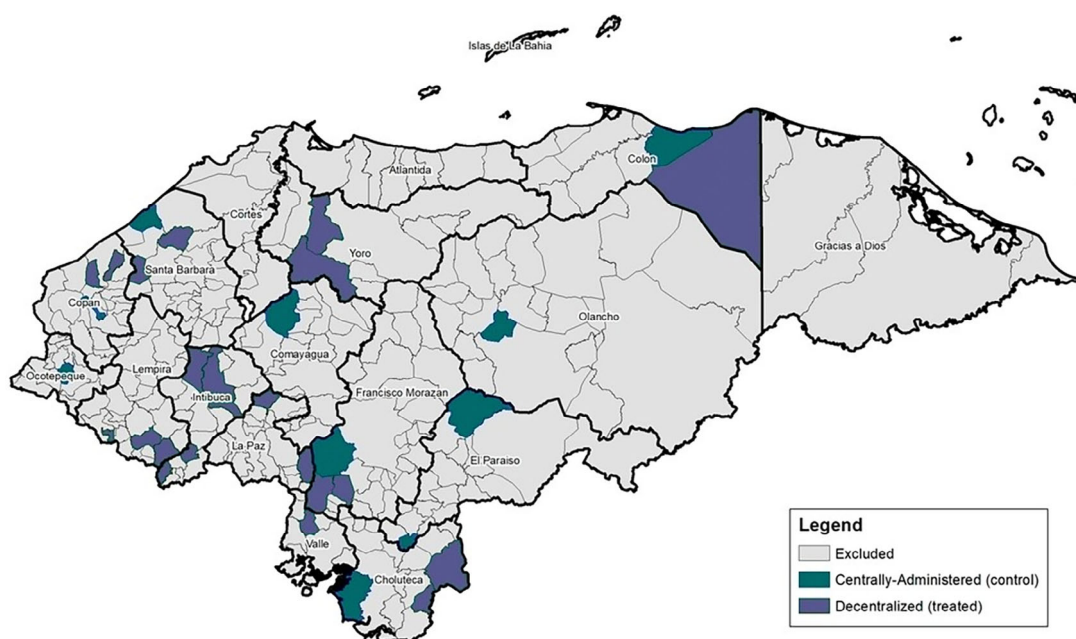


FIGURE 2 Sample of 31 municipalities. An earlier version of this figure first appeared in Zarychta et al. (2019)

propensity decentralized municipalities with relatively weaker matches in the control group. So, we also identified a reduced sample of 24 municipalities by excluding those higher-propensity cases, which we use as a robustness check where there is better balance between the two groups in terms of overall propensity for targeting by the decentralization reform. We generally favor the full sample within the main text of this article due to a stark tradeoff between balance and statistical power. However, we include an analysis using the reduced subsample as an additional robustness check of our findings (Appendix F), and the results are consistent with those we present in the main text.

4.2 | Study-workshop

Between May and September of 2017, subjects were invited to attend a “study-workshop” in which they would simultaneously participate in a research study and also receive training on strategies for effective leadership, communication, and teamwork in health services administration. Thus, the workshops should have been viewed as relevant to the participants' daily work. In each municipality, we recruited eight officials working at various levels within the health sector. The National Ministry of Health sent letters of support to help us recruit subjects, giving permission for subjects to use work hours where necessary to participate in the workshop. We made several accommodations to ensure that those recruited to participate were able to do so, including providing transportation and changing workshop dates where necessary. As a result, there were not many cases where the individuals designated and recruited were unable to attend.

While participation varied slightly by municipality, the typical composition of subjects included four workers from local community health centers, two representatives from the managing organization (or the equivalent role in the centrally administered setting), one health-related official from the mayor's office, and one representative from the regional health authority. Descriptive statistics (Appendix A) show that about 70% of our participants were women. Overall, we achieved a systematic sample of 232 individuals working in the public health sector. They averaged 37 years of age with 8–9 years of experience working in the health sector. Participants from the decentralized setting were, on average, slightly younger, more often male, and had experience of about half as much time working in the health sector. These differences are largely driven by individuals hired after the reform was already in place, with fewer differences apparent when comparing only those hired in the prereform period (Appendix B). Our participants also look fairly similar with regard to these observable characteristics as the broader population of health center personnel in Honduras based on a survey of health center workers across a larger set of municipalities (Appendix C).

The study-workshop included a series of five behavioral games intended to measure behavioral effects of the reform. The games provided an opportunity to earn some modest compensation, which was provided in kind at the end of the day through a choice of commonly used goods, commensurate with the amount earned. Participants also filled out surveys. We describe the results of these other games elsewhere, but here, we focus exclusively on the puzzle game as it is the only game that measures the outcome of interest in this paper: teamwork.

4.3 | Cross-level teams of civil servants

Our unit of analysis for this game is a two-person team of civil servants from different levels of the public health sector hierarchy. Pairs were not randomly assigned but rather were organized to maximize the number of different types of pairs in each municipality based on the participants' roles in the health service delivery system. For example, a typical team could be comprised of one representative from the Regional Health Authority and a lead doctor from one of the rural health centers supervised by that representative. Thus, the teams are comprised of individuals who would need to coordinate regularly in order for the local health system to achieve its goals but not necessarily individuals who are physically in the same office or health center as each other on a daily basis.¹ The goal of this research design was to bring together all those actors from across the health sector structure that could and should

plausibly interact with each other in coproducing health services and affecting health outcomes at the local level given that this vertical administrative structure was directly altered by the reform under study here.

4.4 | The puzzle game

To measure teamwork, we invited these pairs of subjects to play a puzzle game modeled after prior research on a similar mechanism to explain local public good provision (Habyarimana et al., 2007). In our version, two participants are paired, and each receives one half of a simple, eight-piece wooden jigsaw puzzle. Participants sit on either side of a divider, such that each can only see their own half of the puzzle, not the pieces held by their teammate. Using verbal communication, they must work together to align the puzzle such that player one's pieces will interlock correctly with player two's pieces when the divider is removed. At the end of the game, it is possible for a team to have any number of pairs of puzzle pieces correctly matched, ranging from 0 to 4. This game is a “complex task” because it involves multiple, mutually exclusive paths to the goal (Campbell, 1988).

Players engaged in four rounds of this game, each time with a new partner and a new puzzle. In each round, teams could earn a reward of 50 Lempiras (about \$2 USD) by being the first to successfully solve the puzzle. Local facilitators observed teammate interactions and coded them for various characteristics, including whether either team member took on a leadership role during game play and whether or not the team members used disapproving language with each other. For each team in each round, we recorded the number of pieces that they had correctly aligned when time ran out. This number ranges from 0 to 4 and is our key measure of success in the puzzle game.

Jigsaw puzzles have been commonly invoked in prior literature on teamwork. They serve as a metaphor for how teams working to solve problems arrive at a common ground of understanding (McNeese, 2003), with individual ways of thinking “fitting together like puzzle pieces” (Rentsch et al., 2008, p. 145). Puzzles have a long history of use to simulate team problem-solving tasks (Brandler, 1988) and are used for studying problem solving in teams because they are “an analytical task that requires all team members' contributions” (Qiu et al., 2009, p. 47).

Our game goes further than a traditional jigsaw puzzle in representing effective team problem solving because it is structured to ensure that is impossible for an individual team member to successfully complete the task without involving their teammate. While individual participants could easily put the puzzle together alone with full access to all pieces (see Figure 3 for an example puzzle), in our puzzle game, each teammate has access to only half the pieces. Successful completion of the puzzle game thus depends on the effective communication of information available to the pair as a collective unit but not to either individual alone, as well as coordination of actions by both members.

Thus, this incentivized task is a test not of individual skill but of the ability to communicate and coordinate successfully toward joint problem solving with other health sector workers in one's municipality. As discussed earlier, communication and coordination are essential components of effective teamwork, and our game is a direct measure of

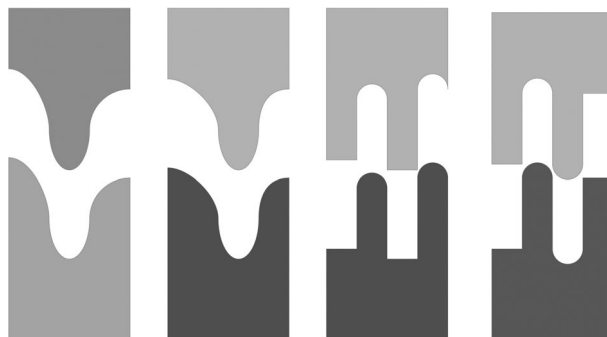


FIGURE 3 Example round of the puzzle game

these team-level skills. The original application of the puzzle game describes it as a measure of whether certain people “work better together,” and a key feature of the game is that “effective communication is a critical determinant of success” (Habyarimana et al., 2007, p. 718).

4.5 | Analytic methods

Our unit of analysis was the two-person team assigned to work on the puzzle game in each round. Our main dependent variable was the number of pairs of puzzle pieces (of four) that a team had correctly aligned when time ran out. The administrative setting (decentralized municipality vs. a centrally administered one) was our key explanatory variable. We also included round/puzzle fixed effects and demographic controls such as gender, age, education, and time working in the health sector. Additional controls included whether teammates knew each other before the workshop, their respective roles in the health sector, self-stated understanding of the game rules, and whether they do puzzles in their spare time as a measure of ability and interest.

Our main models use linear regression with robust standard errors clustered by municipality. Because our unit of analysis is a team of two, our control variables take the form of team-level summaries: the mean of each variable across the two teammates, as well as the difference between them on each variable. All regressions also include the municipal propensity score weights described earlier. As the propensity score matching was based on municipal-level characteristics, it is appropriate to include both the propensity score weights and the team-level demographic characteristics (based on information about the particular individuals who participated in our workshop) within the same regression model.

Because our study workshop included a total of five games, there is some rationale for including corrections for multiple hypothesis testing in our results as a way to reduce the likelihood of making a Type I error (Moore & McCabe, 1999; Rosenthal & Rosnow, 1984). This practice remains controversial, however, as it simultaneously increases the likelihood of a Type II error (Cabin & Mitchell, 2000), and other scholars argue that the practice can lead to unreasonable conclusions if consistently applied (O’Keefe, 2003). In addition, many scholars criticize the reliance on statistical significance testing at all as opposed to, for example, emphasizing the magnitude of the observed effect (Carver, 1978, 1993). For this analysis, we choose to simply note that Bonferroni-corrected significance tests would involve dividing alpha by 5. Thus, if the reader considers $\alpha = 0.05$ to be the appropriate cutoff for unadjusted tests, the adjusted test will use $\alpha = 0.01$ as the cutoff instead. We denote both significance levels in all of our tables.

To test for heterogeneous treatment effects by hiring date, we created an indicator for whether each team member was hired before or after implementation of the reform. (For centrally administered municipalities, we identify the year in which the nearest reform municipality was decentralized and use that to create a comparable indicator of new vs. existing staff.) The team-level aggregate of this indicator takes on a value of 0 if both teammates are new (hired after the reform was already in place), 1 if only one of the two teammates is new, or 2 if both teammates are existing staff (already working in the health sector before the reform began). We include an interaction effect between this new variable and the reform treatment, allowing us to compare the effect of the reform on teams of varying composition. Finally, we examine leadership and communication as potential mediators of how the reform affects behavioral outcomes.

5 | RESULTS

5.1 | Similar teams are less effective at working together in the decentralized setting

Table 1 reports the effect of the decentralization reform on successful teamwork, conditional on several team-level control variables (Appendix E). Our main models suggest the reform had a negative effect on teamwork. On average,

TABLE 1 Effect of governance reform on successful teamwork

	(1)	(2)	(3)	(4)	(5)	(6)
Decentralization	0.0967 (0.193)	−0.368* (0.154)	−0.488** (0.159)	−0.535** (0.163)	−0.539** (0.164)	−0.592*** (0.155)
Team-level controls included:						
Gender		Y	Y	Y	Y	Y
Age		Y	Y	Y	Y	Y
Education			Y	Y	Y	Y
Years of experience			Y	Y	Y	Y
Puzzles free time				Y	Y	Y
Understood game				Y	Y	Y
Knows partner					Y	Y
Dyad-type fixed effects						Y
Round fixed effects	Y	Y	Y	Y	Y	Y
R ²	0.056	0.116	0.141	0.171	0.172	0.189
Adjusted R ²	0.047	0.100	0.115	0.137	0.134	0.138
Observations	429	425	411	408	408	408

Note: Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

after controlling for demographic factors, teams in decentralized settings complete 0.59 fewer puzzle pieces than their counterparts under centralized administration ($p = 0.001$), representing a 15% difference in the total proportion completed. The negative coefficient is larger in magnitude than for any control variables.

This main result holds under various robustness checks, including alternate model specifications; the inclusion of department fixed effects; and using a reduced subsample that improves balance between treated and control municipalities (Appendix F), as well as alternate specifications of the outcome variable (Appendix G). Breaking down the results by type of managing organization shows consistent results as well—the coefficient for all three types is negative and is also statistically significant (except in the case of municipalities, for which we had a slightly smaller sample size).² Not all alternate models yield statistical significance at the smaller Bonferroni-corrected alpha level, but the negative direction and approximate magnitude of the result is consistent.

5.2 | Less effective teamwork under decentralization is driven by staff hired prior to reform

Figure 4 shows our analysis of heterogeneous treatment effects (full models in Appendix H). We find that the negative effect of the reform disappears when the team is made up only of new staff members; the coefficient is positive but lacks statistical significance. Evidence for the negative effect is strongest among mixed groups that have one new and one existing staff member working together, but even among teams comprised only of existing staff, the negative effect persists. Removing the demographic controls yields a similar pattern except that it is the positive effect on new staff that is statistically significant when we do not control for differences in composition.

These results suggest there is a positive effect on new staff that operates through a selection mechanism: new staff hired under the reform appear to be able to coordinate better on this kind of complex task, at least with each other. This is consistent with previous research on selection effects—new staff hired under the reform tended to be more motivated (Zarychta, Grillos & Andersson, 2019). In contrast, the negative effect on existing staff is related to

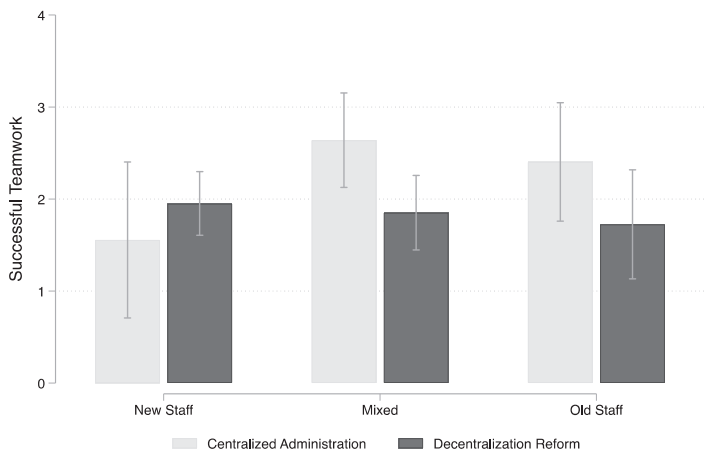


FIGURE 4 Predicted team success by reform setting and hiring date

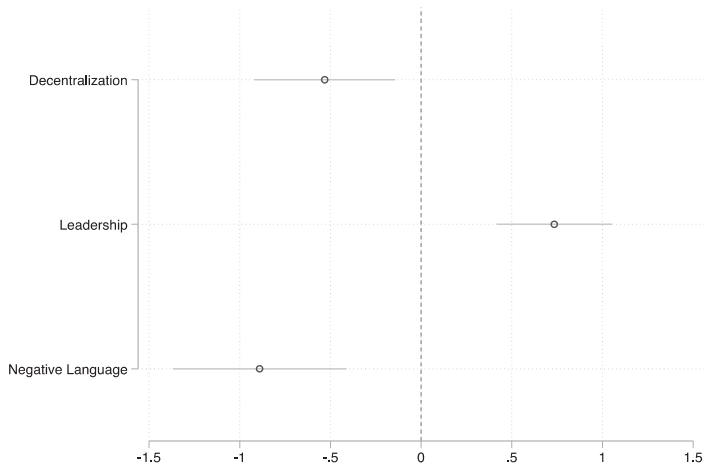


FIGURE 5 Predictors of successful teamwork

socialization: the same kinds of people, confronted with this change to their daily work environment, had greater difficulty in coordinating with one another on a joint problem-solving task that requires effective communication.

5.3 | Negative language partially mediates the adverse effects of decentralization on teamwork

Negative language was relatively rare during the course of the puzzle activity, but it occurred about twice as often in the decentralized setting (10.96% vs. only 5.11% of the time). This correlation retains marginal statistical significance ($p = 0.089$) even after adding the full set of model controls (Appendix J). The average causal mediation effect (Hicks & Tingley, 2011) of negative language on successful teamwork is -0.06 (95% confidence interval: $-0.116, -0.009$), accounting for around 11% of the total effect. Sensitivity analysis demonstrates that the effect remains positive as long as the negative correlation between error terms in the two equations (predicting the mediator and the outcome, respectively) is no greater than 0.3.

5.4 | Leadership is strongly associated with effective teamwork

Figure 5 shows the effects of the decentralization reform on teamwork, as well as the effects of leadership and the use of negative language. Consistent with recent findings on voluntary leadership (b; Andersson et al., 2020a), when observers recorded that one player chose to take a leadership role in the game, this increased completion by about three-quarters of a puzzle piece. The reform was not a predictor of leadership, and the effect of leadership on outcomes is robust to controlling for the reform setting. On average, the positive effect of leadership on effective teamwork is larger than the negative consequences of the decentralization reform (Appendix I).

6 | LIMITATIONS

Although our work was conducted in partnership with the Honduran Ministry of Health, the reform was already underway when we collected our data, and so, we could not influence the selection of municipalities for early rollout. Assignment to the treatment group is not random, and we cannot make the same claim of causal inference that an experimental approach could (Holland, 1986; Sekhon, 2009). However, even some randomized trials fail to achieve true randomization in practice, and propensity score matching is an appropriate way to estimate causal effects in such cases (Imai, 2005). Our matching procedure accounts for observable differences between treatment and control communities on all variables that theoretically influence outcomes. That said, we cannot completely rule out the possibility of additional unobservable differences between the groups that theoretical literature has not yet identified.

We also rely on a simulated behavioral game to measure our primary outcome of interest, teamwork. Our goal was to measure teamwork across various levels and role types in the health sector. Thus, in the daily work environment, each application of teamwork would have a different goal and a different way of measuring success. This would make observations of teamwork in the field both extremely difficult to implement across the number of municipalities in our sample and also challenging to directly compare across settings. The games approach mitigates these issues by providing more controlled measurement, but it does raise other concerns about the extent to which the game truly reflects teamwork in more contextualized settings (Grossman, 2011). While some prior studies have shown correlations between real-world behavior and behavior within games designed to simulate those situations (Benz & Meier, 2008; Grossman & Baldassarri, 2012), ultimately, this study's results hinge on the degree to which real-world teamwork requires the same critical components necessary to achieve success in our game: effective communication and coordination with others who have different access to knowledge and resources.

7 | DISCUSSION AND CONCLUSION

We find that cross-level teams of civil servants in municipalities under decentralization reform are less effective at a joint problem-solving task compared with otherwise comparable teams in the centrally administered setting. This effect persists despite selection effects that would directly counteract it. After experiencing the reform, teams with preexisting staff are less successful at working together in a timed problem-solving challenge that requires knowledge and input from both parties.

This raises questions about the effectiveness of contemporary decentralization reforms with respect to performance within the health sector, which at least in some cases requires effective communication and coordination across work units. The implication for particular policy outcomes depends on the kinds of tasks that bureaucrats most frequently engage in to achieve those outcomes. For relatively routine, simple tasks, such as postconsultation paperwork, the kind of teamwork measured here may be inconsequential. However, for novel, complex problem-solving tasks, our results are worrisome. Tricky diagnoses that involve more than one team member or determining

the correct supply order under changing circumstances may fall into this category and thus be affected by the lack of coordination measured in our game. Health sector workers may face both types of tasks in their daily work, and expanding this typology of bureaucratic actions may be a fruitful area for future research.

More generally, our findings point to the potential disruption that policy change can wreak on the effectiveness of existing staff members, at least with respect to complex tasks. The negative results we report are driven by staff members, hired before the reform, who experienced a change to their work environment. This effect was large enough to outweigh countervailing positive effects related to the hiring of new staff. New ways of doing things, even if they provide a net benefit overall, likely have transaction or adjustment costs among existing public employees with corresponding behavioral impacts on their ability to communicate and work effectively as a team. Future qualitative research could try to unpack specific mechanisms through which the reform may have engendered negative reactions from existing staff—general uncertainty, resentment toward new staff, specific concerns with the reform, or other factors.

Given the importance of institutional knowledge and value of hands-on experience, it is desirable to retain and make use of strengths within the existing workforce of civil servants. Would-be reformers should therefore place priority on the need to help existing staff members transition to a substantially altered work environment. Our results also point to the potential importance of leadership in aiding that transition. This finding complements previous scholarship focusing on the importance of leaders in “setting the table” in terms of establishing a positive organizational culture (Moynihan et al., 2011), suggesting that offering more opportunities for leadership training in local teams could help overcome the noted behavioral challenges associated with contemporary governance reforms. This may be particularly important in sectors like health, education, and environmental management where service delivery often requires the resolution of novel, complex problems by teams of civil servants.

Overall, the ability of civil servants to work together effectively is an integral component of the administrative capacity needed to deliver public services to local communities, especially where complex tasks are common. Policymakers seeking to pursue governance reform must contend with the need to help existing staff members as they transition to a new and different work environment.

ACKNOWLEDGEMENTS

This research was conducted with support from the National Science Foundation (Award numbers SMA-1328688 and DGE-1144083). We wish to thank the Honduran Ministry of Health for its support of this project, particularly the Undersecretary of Integrated Health Service Networks and the Unit for Decentralized Management, as well as Adriana Molina-Garzón and our team of Honduran enumerators for their assistance with fieldwork, data collection and entry. Jerry Patrasso helped us design and produce the wooden puzzle pieces that we used in the experiment. We also wish to thank Jane Menken, Elisabeth Root and Bertha Bermudez Tapia for contributions to the broader research program that supported this analysis.

ENDNOTES

¹ This is in line with the importance the World Health Organization (WHO) places on interprofessional healthcare teams, which is when “...multiple health workers from different professional backgrounds provide comprehensive services by working with patients, their families, carers and communities to deliver the highest quality of care across settings,” and the concept of a health system that consists of “...all the organizations, people and actions whose primary intent is to promote, restore or maintain health...” (WHO, 2010, p. 13).

² In other work (Root et al., 2020), we find that the type of managing organization does matter for health outcomes, with municipalities led by Non-Governmental Organizations (NGOs) performing the best. For the analysis presented in this paper, a direct comparison of the treatment effects for each managing organization yields a statistically significant difference only when comparing NGOs to associations of mayors, for which the coefficients are -0.48 and -82 , respectively.

DATA AVAILABILITY STATEMENT

Data available on request from the authors.

ORCID

Tara Grillos  <https://orcid.org/0000-0001-5743-3825>

Alan Zarychta  <https://orcid.org/0000-0002-4342-2356>

Krister Andersson  <https://orcid.org/0000-0002-9320-8155>

REFERENCES

- Andersson, K. (2004) Who talks with whom? The role of repeated interactions in decentralized Forest governance. *World Development*, 32(2), 233–249.
- Andersson, K., Gordillo, G. & van Laerhoven, F. (2009) *Local governments and rural development: comparing lessons from Brazil, Chile, Mexico, and Peru*. Tucson, AZ: University of Arizona Press.
- Andersson, K.P., Chang, K. & Molina-Garzón, A. (2020a) Voluntary leadership and the emergence of institutions for self-governance. *Proceedings of the National Academy of Sciences*, 117, 27292–27299.
- Andersson, K.P., Chang, K. & Molina-Garzón, A. (2020b) Voluntary leadership and the emergence of institutions for self-governance. *Proceedings of the National Academy of Sciences*, 117(44), 27063–27071.
- Ang, Y.Y. (2017) Beyond weber: conceptualizing an alternative ideal type of bureaucracy in developing contexts. *Regulation & Governance*, 11(3), 282–298.
- Ansell, C. & Gash, A. (2018) Collaborative platforms as a governance strategy. *Journal of Public Administration Research and Theory*, 28(1), 16–32.
- Arellano-Gault, D. (2000) Challenges for the new public management: organizational culture and the administrative modernization program in Mexico City. *American Review of Public Administration*, 30(4), 400–413.
- Arellano-Gault, D. & Ramírez, E. (2008) Performance-based budgeting in Latin and South America. In: *International handbook of practice-based performance management*. Thousand Oaks, California: Sage Publications.
- Ateş, H. (2004) Management as an agent of cultural change in the Turkish public sector. *Journal of Public Administration Research and Theory*, 14(1), 33–58.
- Avellaneda, C.N. & Gomes, R.C. (2015) Is small beautiful? Testing the direct and nonlinear effects of size on municipal performance. *Public Administration Review*, 75(1), 137–149.
- Baker, D.P., Day, R. & Salas, E. (2006) Teamwork as an essential component of high-reliability organizations. *Health Services Research*, 41(4), 1576–1598.
- Baldwin, D.C. & Daugherty, S.R. (2008) Interprofessional conflict and medical errors: results of a national multi-specialty survey of hospital residents in the US. *Journal of Interprofessional Care*, 22(6), 573–586.
- Batley, R. (1999) The new public management in developing countries: implications for policy and organizational reform. *Journal of International Development*, 11(5), 761–765.
- Battaglio, R.P., Jr., Belardinelli, P., Bellé, N. & Cantarelli, P. (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), 304–320.
- Benz, M. & Meier, S. (2008) Do people behave in experiments as in the field? Evidence from donations. *Experimental Economics*, 11(3), 268281281.
- Bhanot, S.P. & Linos, E. (2020) Behavioral public administration: past, present, and future. *Public Administration Review*, 80(1), 168–171.
- Blair, H. (2000) Participation and accountability at the periphery: democratic local governance in six countries. *World Development*, 28(1), 21–39.
- Boschi, R. & Gaitán, F. (2009) Legados, política y consenso desarrollista. *Nueva Sociedad*, 224, 14.
- Bossert, T., Chitah, M.B. & Bowser, D. (2003) Decentralization in Zambia: resource allocation and district performance. *Health Policy and Planning*, 18(4), 357–369.
- Bowles, S. (2008) Policies designed for self-interested citizens may undermine "the moral sentiments": evidence from economic experiments. *Science*, 320(5883), 1605–1609.
- Brandler, S. (1988) The jigsaw puzzle: an experiment in understanding group process. *Social Work with Groups*, 11(1–2), 99–109.
- Brehm, J.O. & Gates, S. (1999) *Working, shirking, and sabotage: bureaucratic response to a democratic public*. Ann Arbor: University of Michigan Press.
- Brinkerhoff, D.W. & Brinkerhoff, J.M. (2015) Public sector management reform in developing countries: perspectives beyond NPM orthodoxy. *Public Administration and Development*, 35(4), 222–237.
- Brodkin, E.Z. (1990) Implementation as policy politics. In: Palumbo, D.J. & Calista, D.J. (Eds.) *Implementation as the policy process: opening up the black box*. Westport, CT: Greenwood Press, pp. 107–118.
- Brodkin, E.Z. (2008) Accountability in street-level organizations. *International Journal of Public Administration*, 31(3), 317–336.
- Brodkin, E.Z. (2011) Policy work: street-level organizations under new managerialism. *Journal of Public Administration Research and Theory*, 21(suppl_2), i253–i277.

- Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., et al. (2013) Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365–376.
- Cabin, R.J. & Mitchell, R.J. (2000) To Bonferroni or not to Bonferroni: when and how are the questions. *Bulletin of the Ecological Society of America*, 81(3), 246–248.
- Campbell, D.J. (1988) Task complexity: a review and analysis. *Academy of Management Review*, 13(1), 40–52.
- Campos, J. & Hellman, J. (2005) Governance gone local: does decentralization improve accountability. In: *East Asia decentralizes*. Washington DC: World Bank.
- Cárdenas, S. & Ramirez de la Cruz, E.E. (2017) Controlling administrative discretion promotes social equity? Evidence from a natural experiment. *Public Administration Review*, 77(1), 80–89.
- Carmenate-Milián, L., et al. (2017) Situation of the health system in Honduras and the new proposed health model. *Archives of Medicine*, 9(4), 1–8.
- Carver, R. (1978) The case against statistical significance testing. *Harvard Educational Review*, 48(3), 378–399.
- Carver, R.P. (1993) The case against statistical significance testing, revisited. *The Journal of Experimental Education*, 61(4), 287–292.
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1), 155–159.
- Danjoux Meth, N., Lawless, B. & Hawryluck, L. (2009) Conflicts in the ICU: perspectives of administrators and clinicians. *Intensive Care Medicine*, 35(12), 2068–2077.
- Deci, E.L., Koestner, R. & Ryan, R.M. (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668.
- Diamond, L. & Tsalkis, S. (1999) Size and democracy: the case for decentralization. In: *Developing democracy: toward consolidation*. Baltimore: Johns Hopkins University Press.
- Dixit, A. (2002) Incentives and organizations in the public sector: an interpretative review. *Journal of Human Resources*, 37(4), 696–727.
- Dunning, T. (2012) *Natural experiments in the social sciences: a design-based approach*. Cambridge: Cambridge University Press.
- El Bcheraoui, C., et al. (2017) Healthy competition drives success in results-based aid: lessons from the Salud Mesoamérica initiative. *PLoS One*, 12(10), e0187107.
- Emerson, K., Nabatchi, T. & Balogh, S. (2012) An integrative framework for collaborative governance. *Journal of Public Administration Research and Theory*, 22(1), 1–29.
- Faguet, J.P. (2004) Does decentralization increase government responsiveness to local needs? Evidence from Bolivia. *Journal of Public Economics*, 88, 667–693.
- Faguet, J.P. (2012). *Decentralization and popular democracy: Governance from below in Bolivia*. Ann Arbor: University of Michigan Press.
- Faguet, J.P. (2014) Decentralization and governance. *World Development*, 53, 2–13.
- Finn, R., Learmonth, M. & Reedy, P. (2010) Some unintended effects of teamwork in healthcare. *Social Science & Medicine*, 70, 1148–1154.
- Frey, B.S. (1994) How intrinsic motivation is crowded out and in. *Rationality and Society*, 6(3), 334–352.
- Fussell S.R., Kraut R.E., Lerch F.J., Scherlis W.L., McNally M.M. & Cadiz J.J. (1998) Coordination, overload and team performance: effects of team communication strategies. *Proceedings of CSCW*, 1998, 275–284.
- Gaitán, F. (2014) *Auge, ocaso y resurgimiento de los estudios sobre desarrollo en América Latina*. Documento de Proyecto, CEPAL. Santiago de Chile: CEPAL.
- Galiani, S., Gertler, P. & Schargrodsky, E. (2008) School decentralization: helping the good get better, but leaving the poor behind. *Journal of Public Economics*, 92(10–11), 2106–2120 <https://doi.org/10.1016/j.jpubeco.2008.05.004>.
- García Prado, A. & Peña, C.L. (2010) *Contracting and providing basic health Care Services in Honduras: a comparison of traditional and alternative service delivery models*. Washington, DC: World Bank. Health Nutrition and Population Discussion Paper.
- Geddes, B. (1994) *Politician's dilemma: building state capacity in Latin America*. Berkeley: University of California Press.
- Gilson, L. (2003) Trust and the development of health care as a social institution. *Social Science & Medicine*, 56(7), 1453–1468.
- Gneezy, U. & Rustichini, A. (2000) Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115(3), 791–810.
- Grimmelikhuijsen, S., Jilke, S., Olsen, A.L. & Tummers, L. (2016) Behavioral public administration: combining insights from public administration and psychology. *Public Administration Review*, 77(1), 45–56.
- Grossman, G. (2011) Lab-in-the-field experiments. *Newsletter of the APSA Experimental Section*, 2(2), 13–19.
- Grossman, G. & Baldassarri, D. (2012) The impact of elections on cooperation: evidence from a lab-in-the-field experiment in Uganda. *American Journal of Political Science*, 56(4), 964–985.
- Guanais, F.C. & Macinko, J. (2009) The health effects of decentralizing primary care in Brazil. *Health Affairs*, 28(4), 1127–1135.
- Habyarimana, J., Humphreys, M., Posner, D.N. & Weinstein, J.M. (2007) Why does ethnic diversity undermine public goods provision? *American Political Science Review*, 101(4), 709–725.
- Hassan, S. & Wright, B.E. (2019) The behavioral public administration movement: a critical reflection. *Public Administration Review*, 80(1), 163–167.
- Hayek, F.A. (1945) The use of knowledge in society. *The American Economic Review*, 35(4), 519–530.

- Heyman, J. & Ariely, D. (2004) Effort for payment—a tale of two markets. *Psychological Science*, 15, 787–793.
- Hicks, R. & Tingley, D. (2011) Causal mediation analysis. *The Stata Journal*, 11(4), 605–619.
- Holland, P.W. (1986) Statistics and causal inference. *Journal of the American Statistical Association*, 81(396), 945–960.
- Hood, C. (1991) A public management for all seasons? *Public Administration*, 69(1), 3–19.
- Hope, K.R., Sr. & Chikulo, B.C. (2000) Decentralization, the new public management, and the changing role of the public sector in Africa. *Public Management and International Journal of Research and Theory*, 2(1), 25–42.
- Hupe, P. & Hill, M. (2007) Street-level bureaucracy and public accountability. *Public Administration*, 85(2), 279–299.
- Imai, K. (2005) Do get-out-the-vote calls reduce turnout? The importance of statistical methods for field experiments. *American Political Science Review*, 99(2), 283–300.
- Imbens, G.W. & Rubin, D.B. (2015) *Causal inference for statistics, social, and biomedical sciences*. New York, NY: Cambridge University Press.
- Jacobsen, C.B., Hvitved, J. & Andersen, L.B. (2014) Command and motivation: how the perception of external interventions relates to intrinsic motivation and public service motivation. *Public Administration*, 92, 790–806.
- Jakobsen, M.L.F. & Mortensen, P.B. (2016) Rules and the doctrine of performance management. *Public Administration Review*, 76, 302–312.
- Jakobsen, M.L., Baekgaard, M., Moynihan, D.P. & van Loon, N. (2018) Making sense of performance regimes: rebalancing external accountability and internal learning. *Perspectives on Public Management and Governance*, 1(2), 127–141.
- Jeppsson, A. & Okunzi, S.A. (2000) Vertical or holistic decentralization of the health sector? Experiences from Zambia and Uganda. *The International Journal of Health Planning and Management*, 15(4), 273–289.
- Jiang, J. (2018) Making bureaucracy work: patronage networks, performance incentives, and economic development in China. *American Journal of Political Science*, 62(4), 982–999.
- Johnson, C. (2009) Bad blood: doctor-nurse behavior problems impact patient care. *Physician Executive Journal*, 35(6), 6–11.
- Langbein, L. & Sanabria, P. (2017) Independent professional bureaucracies and street-level bribery: comparing changes in civil service law and implementation in Latin America. *Journal of Comparative Policy Analysis: Research and Practice*, 19(5), 435–451.
- Larbi, G. (2001) Performance contracting in practice: experience and lessons from the water sector in Ghana. *Public Management Review*, 3(3), 305–324.
- Le Grand, J. (2010) Knights and knaves return: public service motivation and the delivery of public services. *International Public Management Journal*, 13(1), 56–71.
- Leggat, S.G. (2007) Effective healthcare teams require effective team members: defining teamwork competencies. *BMC Health Services Research*, 7(1), 17.
- Lipsky, M. (1980) *Street-level bureaucracy: dilemmas of the individual in public service*. New York: Russell Sage Foundation.
- Loevinsohn, B. & Harding, A. (2005) Buying results? Contracting for health service delivery in developing countries. *The Lancet*, 366(9486), 676–681.
- Magrath, P. & Nichter, M. (2012) Paying for performance and the social relations of health care provision: an anthropological perspective. *Social Science & Medicine*, 75(10), 1778–1785.
- Manning, N. (2001) The legacy of the new public Management in Developing Countries. *International Review of Administrative Sciences*, 67(2), 297–312.
- Marks, M.A., Mathieu, J.E. & Zaccaro, S.J. (2001) A temporally based framework and taxonomy of team processes. *The Academy of Management Review*, 26(3), 356–376.
- May, P.J. & Winter, S.C. (2009) Politicians, managers, and street-level bureaucrats: influences on policy implementation. *Journal of Public Administration Research and Theory*, 19(3), 453–476.
- Maynard-Moody, S.W., Musheno, M. & Musheno, M.C. (2003) *Cops, teachers, counselors: stories from the front lines of public service*. Ann Arbor: University of Michigan Press.
- McNease, M.D. (2003) Metaphors and paradigms of team cognition: a twenty year perspective. In: *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 47, No. 3. Los Angeles, CA: SAGE Publications, pp. 518–522.
- Meyers, M.K., Glaser, B. & Donald, K.M. (1998) On the front lines of welfare delivery: are workers implementing policy reforms? *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 17(1), 1–22.
- Miller, G.J. (1992) *Managerial dilemmas: the political economy of hierarchy*. Cambridge: Cambridge University Press.
- Ministry of Health (MOH), Government of Honduras. (2009) Marco Conceptual Político y Estratégico de La Reforma Del Sector de Salud.
- Ministry of Health (MOH), Government of Honduras. (2010) Plan Nacional de Salud 2010–2014.
- Ministry of Health (MOH), Government of Honduras, Unit for Decentralized Management. (2018) *Guía de Monitoreo Para Primer Nivel de Gestion Descentralizada: Indicadores Seleccionados Para UAPS Adn CIS Con Gestion Por Resultados*. Honduras, CA: Tegucigalpa.
- Moore, D.S. & McCabe, G.P. (1999) *Introduction to the practice of statistics*, 3rd edition. New York: W. H. Freeman.

- 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.
- Moynihan, D.P. & Ingraham, P.W. (2003) Look for the silver lining: when performance-based accountability systems work. *Journal of Public Administration Research and Theory*, 13(4), 469–490.
- Moynihan, D.P. & Soss, J. (2014) Policy feedback and the politics of administration. *Public Administration Review*, 74(3), 320–332.
- Moynihan, D.P., Pandey, S.K. & Wright, B.E. (2011) Setting the table: how transformational leadership fosters performance information use. *Journal of Public Administration Research and Theory*, 22(1), 143–164.
- Nichter, M.A. (1986) The primary health center as a social system: PHC, social status, and the issue of team-work in South Asia. *Social Science & Medicine*, 23(4), 347–355.
- Oates, W.E. (1977) *The political economy of fiscal federalism*, Lexington, MA: Lexington Books.
- O'Keefe, D.J. (2003) Colloquy: should familywise alpha be adjusted? Against familywise alpha adjustment. *Human Communication Research*, 29(3), 431–447.
- Panizza, F. (2004) A reform without loser: the symbolic effects of civil service reform in Uruguay 1995–96. *Latin American Politics and Society*, 46(3), 1–28.
- Panizza, F. & Philip, G. (2005) Second generation reform in Latin America: reforming the public sector in Uruguay and Mexico. *Journal of Latin America Studies*, 37, 667–691.
- Piña, G. & Avellaneda, C.N. (2018) Municipal isomorphism: testing the effects of vertical and horizontal collaboration. *Public Management Review*, 20(4), 445–468.
- Pollitt, C., Van Thiel, S. & Homburg, V. (2007) New Public Management in Europe, Management Online Review.
- Powell-Jackson, T., Morrison, J., Tiwari, S., Neupane, B.D. & Costello, A.M. (2009) The experiences of districts in implementing a national incentive programme to promote safe delivery in Nepal. *BMC Health Services Research*, 9(1), 97.
- Pressman, J.L. & Wildavsky, A. (1979) *Implementation: how great expectations in Washington are dashed in Oakland*, Berkeley: Univ of California Press.
- Qiu, L., Tay, W.W. & Wu, J. (2009) The impact of virtual teamwork on real-world collaboration. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, pp. 44–51.
- Ramos, C. & Milanesi, A. (2020) The neo-Weberian state and the neodevelopmentalist strategies in Latin America: the case of Uruguay. *International Review of Administrative Sciences*, 86(2), 261–277.
- Raudenbush, S.W., Spybrook, J., Congdon, R., Liu, X.F., Martinez, A., Bloom, H. & Hill, C. (2011) Optimal design software for multi-level and longitudinal research (Version 3.01) [Software].
- Rentsch, J.R., Small, E.E. & Hanges, P.J. (2008) Cognitions in organizations and teams: what is the meaning of cognitive similarity? In: Smith, B. (Ed.) *The people make the place: exploring dynamic linkages between individuals and organizations*. New York, NY: Erlbaum.
- Ribot, J.C. (2002) *Democratic decentralization of natural resources: institutionalizing popular participation*. Washington, DC: World Resources Institute.
- Riccucci, N.M., Meyers, M.K., Lurie, I. & Han, J.S. (2004) The implementation of welfare reform policy: the role of public managers in front-line practices. *Public Administration Review*, 64(4), 438–448.
- Rigoli, F. & Dussault, G. (2003) The interface between health sector reform and human resources in health. *Human Resources for Health*, 1(1), 9.
- Robalino, D.A., Picazo, O.F. & Voetberg, A. (2001) *Does fiscal decentralization improve health outcomes? Evidence from a cross-country analysis*, Policy Research Working Paper, No. 2565 Washington DC: World Bank.
- Rondinelli, D.A., McCullough, J.S. & Johnson, R.W. (1989) Analysing decentralization policies in developing countries: a political-economy framework. *Development and Change*, 20(1), 57–87.
- Root, E.D., Zarychta, A., Tapia, B.B., Grillos, T., Andersson, K. & Menken, J. (2020) Organizations matter in local governance: evidence from health sector decentralization in Honduras. *Health Policy and Planning*, 35(9), 1168–1179. <https://doi.org/10.1093/heapol/czaa084>.
- Rosenthal, R. & Rosnow, R.L. (1984) *Essentials of behavioral research: methods and data analysis*. New York: McGraw-Hill.
- Ryan, R.M. & Weinstein, N. (2009) Undermining quality teaching and learning: a self-determination theory perspective on high-stakes testing. *Theory and Research in Education*, 7, 224–233.
- Salas, E., Dickinson T.L., Converse S.A. & Tannenbaum S.I. (1992) Toward an understanding of team performance and training. In: Swezey, R.W. & Salas, E. (Eds.) *Teams: their training and performance*. Norwood, NJ: Ablex, pp. 3–29.
- Scholz, J.T., Twombly, J. & Headrick, B. (1991) Street-level political controls over Federal bureaucracy. *American Political Science Review*, 85(3), 829–850.
- Sekhon, J.S. (2009) Opiates for the matches: matching methods for causal inference. *Annual Review of Political Science*, 12, 487–508.
- Solheim, K., McElmurry, B.J. & Kim, M.J. (2007) Multidisciplinary teamwork in US primary health care. *Social Science & Medicine*, 65(3), 622–634.

- Soss, J., Fording, R.C. & Schram, S. (2011) *Disciplining the poor: neoliberal paternalism and the persistent power of race*. Chicago: University of Chicago Press.
- Stepan, A. (2001) Toward a new comparative politics of federalism, (multi)nationalism, and democracy: beyond Rikerian federalism. In: *Alfred Stepan arguing comparative politics*, Oxford: Oxford University Press, pp. 315–361.
- Tankha, S. (2009) Building administrative capacities in developing countries: SWAT teams or beat cops? *Public Administration Review*, 69(6), 1028–1030.
- Weaver, S.J., Feitosa, J., Salas, E., Seddon, R. & Vozenilek, J.A. (2012) The theoretical drivers and models of team performance and effectiveness for patient safety. In: *Improving patient safety through teamwork and team training*, Oxford: Oxford University Press.
- World Bank. (2004) *World development report 2004: making services work for poor people*. Washington, DC: The World Bank & Oxford University Press.
- World Health Organization (WHO). (2010) *Framework for action on Interprofessional education & collaborative practice*. Geneva, Switzerland: World Health Organization, Department of Human Resources for Health. https://www.who.int/hrh/resources/framework_action/en/.
- 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, <https://doi.org/10.1016/j.worlddev.2020.104996>.
- Zarychta, A., Grillos, T., & Andersson, K. (2019). Public Sector Governance Reform and the Motivation of Street-Level Bureaucrats in Developing Countries. *Public Administration Review*, 80(1), 75–91. <https://doi.org/10.1111/puar.13132>.
- Zarychta, A., Andersson, K., Root, E., Menken, J., & Grillos, T. (2019). Assessing the impacts of governance reforms on health services delivery: a systematic approach to overcome persistent weaknesses of policy evaluation. *Health Services and Outcomes Research Methodology*, 19(4), 241–258. <https://doi.org/10.1007/s10742-019-00201-8>.

How to cite this article: Grillos T, Zarychta A, Andersson K. Governance reform, decentralization, and teamwork in public service delivery: Evidence from the Honduran health sector. *Public Admin.* 2021;1–27. <https://doi.org/10.1111/padm.12722>

APPENDIX A: DESCRIPTIVE STATISTICS

Variable	Full sample				Centralized administration				Decentralization reform			
	n	Mean/prop	SD	Min	Max	n	Mean/prop	SD	n	Mean/prop	SD	SD
Female	232	0.69	0.46	0	1	77	0.83	0.38	155	0.62	0.49	0.49
Age	231	36.77	10.12	20	64	76	40.18	11.04	155	35.09	9.22	9.22
Education	228	4.54	0.72	1	5	74	4.50	0.65	154	4.56	0.75	0.75
Years in health	232	8.52	8.03	0	38	77	12.31	9.48	155	6.63	6.44	6.44
Years in current job	232	6.05	6.95	0	38	77	9.20	8.94	155	4.48	5.05	5.05
Administrator	232	0.34	0.47	0	1	77	0.34	0.48	155	0.34	0.47	0.47
Doctor	232	0.27	0.45	0	1	77	0.17	0.38	155	0.32	0.47	0.47
Nurse	232	0.23	0.42	0	1	77	0.32	0.47	155	0.19	0.39	0.39
Community health worker	232	0.03	0.18	0	1	77	0.03	0.16	155	0.04	0.19	0.19
Mayor's office	232	0.11	0.32	0	1	77	0.13	0.34	155	0.10	0.31	0.31
Plays puzzles	231	2.35	1.13	1	5	77	2.13	1.06	154	2.47	1.15	1.15

APPENDIX B: DESCRIPTIVE STATISTICS FOR PREREFORM STAFF ONLY

	Centralized		Decentralized		Difference	Two-sided t-test p Value
	n	Mean	n	Mean		
Female	54	0.83	51	0.69	0.15	0.08
Age	53	43.51	51	41.45	2.06	0.22
Education	51	4.51	51	4.57	−0.06	0.67
Years in health	54	16.43	51	13.92	2.50	0.08
Years in current job	54	11.97	51	8.84	3.13	0.04*
Understood puzzle game	54	2.96	51	3.20	−0.23	0.11
Plays puzzles	54	2.22	51	2.22	0.01	0.97
Administrator	54	0.43	51	0.49	−0.06	0.51
Doctor	54	0.11	51	0.20	−0.08	0.23
Nurse	54	0.35	51	0.24	0.12	0.19
Promotor	54	0.02	51	0.06	−0.04	0.29
Mayor's office	54	0.09	51	0.02	0.07	0.11

* $p < 0.05$.

APPENDIX C: DESCRIPTIVE STATISTICS FROM A NATIONALLY REPRESENTATIVE SURVEY OF HEALTH CENTER PERSONNEL

To address concerns about the representativeness of our sample, below, we present the descriptive statistics from a nationally representative survey of health center workers that were collected as part of the larger National Science Foundation (NSF) grant to which this analysis belongs. Table C1 shows the gender, age, education, years in the health sector, and years working at the current job across the entire sample from the nationally representative survey. Table C2 shows the same descriptive statistics from the same nationally representative survey but for only the subset of municipalities that was selected for inclusion in our games subsample. Finally, Table C3 shows the same descriptive statistics again, this time for the subset of our study workshop participants who are health center workers (as they are the group most directly comparable to who was surveyed in the nationally representative sample). As you can see, our sample is fairly similar to the larger population in terms of gender, age, and years of work experience, but our sample is slightly more likely to have a college degree.

TABLE C1 Descriptive statistics across all 65 municipalities included in nationally representative survey of health center workers ($N = 627$)

	<i>n</i>	Mean/prop.	SD	Min.	Max.
Female	624	0.74	0.44	0.00	1.00
Age	595	35.29	9.62	17.00	74.00
Education	623	3.78	0.77	1.00	5.00
Years in Health	627	8.12	7.69	0.00	37.00
Years in Current Job	627	6.17	6.65	0.00	35.00

TABLE C2 Descriptive statistics from the 31 municipalities in our games sample as they appear in nationally representative survey of health center workers ($N = 359$)

Across all games municipalities					
	<i>n</i>	Mean	SD	Min.	Max.
Female	356	0.71	0.46	0.00	1.00
Age	347	35.05	9.51	17.00	74.00
Education	357	3.73	0.78	2.00	5.00
Years in health	359	7.48	7.01	0.00	36.00
Years in current job	359	5.62	5.99	0.00	33.00

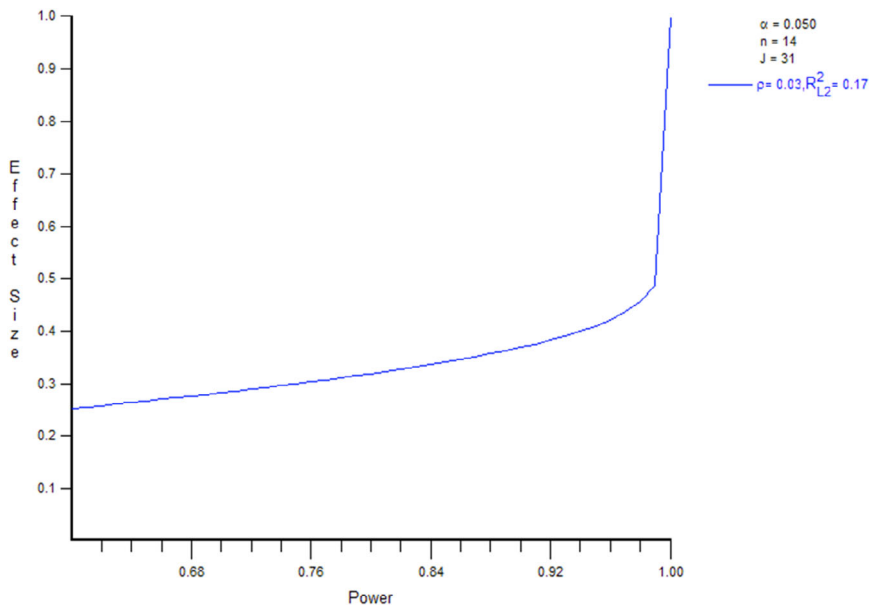
TABLE C3 Descriptive statistics of the health center personnel who participated in our study-workshops ($N = 128$)

	<i>n</i>	Mean/prop.	SD	Min.	Max.
Female	128	0.69	0.47	0.00	1.00
Age	128	34.14	8.70	21.00	55.00
Education	125	4.46	0.69	2.00	5.00
Years in Health	128	7.39	6.90	0.00	30.00
Years in Current Job	128	5.79	6.65	0.00	30.00

APPENDIX D: DISCUSSION OF STATISTICAL POWER

Our analysis included 230 health workers across 31 municipalities. However, puzzle game outcomes were observed at the team (dyad) level, so our sample for this analysis includes 429 distinct two-person combinations (teams), of which 292 (68%) were from the decentralized (treatment) setting. Thus, the main inputs to our power calculations include the following values: $p = 0.68$ (the proportion of the sample that is in the treatment group), $J = 31$ (the number of clusters), and $n = 13.8$ (the average number of teams per cluster). Based on the main model we present in the paper, R^2 is equal to 0.172, and we used Stata's `loneaway` command to estimate the intracluster correlation coefficient, ρ , to be 0.029 for our outcome variable. Assuming standard values for alpha equal to 0.05 and statistical power, $1 - k$, equal to 0.8, our power calculations yield a minimum detectable effect size of 0.31σ . Effect sizes below 0.3σ are considered small (Cohen, 1992), so this suggests that we can detect any moderate effects with a probability of 80%. Using Optimal Design software (Raudenbush et al., 2011), we produced the graph below, which shows estimates of our statistical power to detect different effect sizes. The effect size estimated by our main model is 0.539, which we divide by the standard deviation (1.718) to obtain the standardized effect size of 0.314σ .

The most direct consequence of this level of statistical power is that, if any true effects existed that were smaller than 0.31σ , we would be unlikely to detect them and thus could not draw any strong conclusions from such null results. Given that we do find statistically significant results, the main risk of a small sample size would be its effect



on positive predictive value, PPV (Button et al., 2013), or the probability that an effect observed to be statistically significant is actually true. Following Button et al. (2013), we use the formula $PPV = (Power \times R) / (Power \times R + \alpha)$ and assume a conservative R value of 0.25. At our estimated Power = 0.8 and $\alpha = 0.05$, this yields a PPV of 0.8. At an alpha of 0.01 (a level of significance achieved by our main result in this paper), our statistical power decreases to 60%, and this yields a PPV of 0.94. That means that, based on our assumptions, there is a 94% probability that our observed effect reflects a true effect.

APPENDIX E: MAIN MODEL: FULL REGRESSION OUTPUT

	(1)	(2)	(3)	(4)	(5)	(6)
Decentralized	0.0967 (0.193)	−0.368* (0.154)	−0.488** (0.159)	−0.535** (0.163)	−0.539** (0.164)	−0.592*** (0.155)
Female (Prop.)		−0.716** (0.258)	−0.811** (0.269)	−0.521+ (0.283)	−0.528+ (0.286)	−0.581+ (0.310)
Female (Difference)		0.412* (0.155)	0.402* (0.165)	0.366* (0.173)	0.376* (0.177)	0.358* (0.166)
Age (Mean)		−0.0261* (0.0109)	−0.0229 (0.0180)	−0.0148 (0.0192)	−0.0144 (0.0191)	0.00000240 (0.0194)
Age (Difference)		−0.0197+ (0.00997)	−0.0164 (0.0103)	−0.0127 (0.0102)	−0.0137 (0.0114)	−0.0166 (0.0107)
Education (Mean)			0.474+ (0.271)	0.333 (0.248)	0.325 (0.249)	0.126 (0.286)
Education (Difference)			−0.00251 (0.167)	−0.0645 (0.166)	−0.0724 (0.166)	−0.0795 (0.170)
Years exp. (Mean)			−0.0132 (0.0292)	−0.0135 (0.0302)	−0.0134 (0.0296)	−0.0427 (0.0295)
Years exp. (Difference)			0.00372 (0.0133)	−0.00153 (0.0133)	−0.00184 (0.0134)	0.00358 (0.0135)
Puzzles free Time (Mean)				0.165 (0.107)	0.166 (0.108)	0.139 (0.120)
Puzzles free (Difference)				−0.0380 (0.0951)	−0.0405 (0.0969)	−0.0489 (0.0974)
Understood Game (Mean)				0.517* (0.190)	0.518* (0.190)	0.490* (0.216)
Understood (Difference)				−0.0955 (0.108)	−0.0947 (0.111)	−0.0856 (0.110)
Knows partner (Prop.)					−0.0438 (0.297)	−0.124 (0.304)
Knows partner (Difference)					−0.105 (0.256)	−0.0934 (0.248)
Dyad-Type Fixed Effects (FE)						Y
Round FE	Y	Y	Y	Y	Y	Y
Constant	2.576*** (0.224)	4.408*** (0.460)	2.382 (1.574)	0.676 (1.920)	0.770 (1.924)	1.543 (2.163)
R ²	0.056	0.116	0.141	0.171	0.172	0.189
Adjusted R ²	0.047	0.100	0.115	0.137	0.134	0.138
Observations	429	425	411	408	408	408

Note: Standard errors in parentheses.

+ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

APPENDIX F: ROBUSTNESS CHECKS: MODEL TYPE, REDUCED SAMPLE, AND DEPT. FIXED EFFECTS

	(1)	(2)	(3)	(4)	(5)
Decentralized	−0.725*** (0.208)	−0.586* (0.248)	−0.794* (0.333)	−0.447* (0.196)	−0.550* (0.242)
Female (Mean)	−0.762+ (0.399)	−0.529 (0.321)	−0.696+ (0.407)	−0.255 (0.440)	−0.300 (0.576)
Female (Difference)	0.380+ (0.204)	0.346* (0.167)	0.419* (0.207)	0.542* (0.247)	0.652* (0.304)
Age (Mean)	0.00847 (0.0231)	0.00313 (0.0221)	0.0155 (0.0260)	−0.0224 (0.0190)	−0.0157 (0.0240)
Age (Difference)	−0.0255* (0.0124)	−0.0162 (0.0122)	−0.0256+ (0.0148)	−0.00438 (0.0125)	−0.00843 (0.0142)
Education (Mean)	0.160 (0.352)	0.109 (0.279)	0.0991 (0.337)	0.113 (0.393)	0.0951 (0.427)
Education (Difference)	−0.113 (0.201)	−0.0626 (0.171)	−0.103 (0.200)	0.0629 (0.226)	0.0303 (0.247)
Years of experience (Mean)	−0.0590+ (0.0347)	−0.0485 (0.0345)	−0.0709+ (0.0423)	−0.0164 (0.0307)	−0.0291 (0.0376)
Years of experience (Difference)	0.00630 (0.0157)	0.00318 (0.0138)	0.00732 (0.0163)	0.0142 (0.0174)	0.0171 (0.0210)
Puzzles free time (Mean)	0.131 (0.143)	0.179 (0.139)	0.211 (0.168)	0.127 (0.154)	0.120 (0.185)
Puzzles free time (Difference)	−0.0634 (0.123)	−0.0462 (0.0951)	−0.0658 (0.120)	−0.00648 (0.122)	−0.0217 (0.154)
Understood game (Mean)	0.617* (0.262)	0.460* (0.212)	0.576* (0.257)	0.654** (0.215)	0.739** (0.267)
Understood game (Difference)	−0.118 (0.141)	−0.0795 (0.112)	−0.112 (0.137)	−0.0894 (0.109)	−0.119 (0.145)
Knows partner (Mean)	−0.184 (0.355)	−0.0797 (0.303)	−0.0973 (0.354)	0.143 (0.333)	0.148 (0.386)
Knows partner (Difference)	−0.0675 (0.314)	−0.107 (0.260)	−0.0780 (0.335)	−0.0220 (0.278)	0.0381 (0.332)
Round fixed effects?	Y	Y	Y	Y	Y
Dyad type FE?	Y	Y	Y	Y	Y
Department FE?		Y	Y		
Model type?	meologit	ols	meologit	ols	meologit
Sample?	Full	Full	Full	Reduced	Reduced
Observations	408	408	408	311	311

Note: Standard errors in parentheses.

+ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

APPENDIX G: ROBUSTNESS CHECKS 2: ALTERNATE SPECIFICATIONS OF OUTCOME VARIABLE

	(1) Success (Y/N)	(2) Success (Y/N)	(3) Most pieces ever correct	(4) Most pieces ever correct
Decentralized	−0.116** (0.0418)	−0.591* (0.230)	−0.465** (0.161)	−0.654** (0.218)
Female (Mean)	−0.183* (0.0773)	−0.965* (0.390)	−0.403 (0.318)	−0.680 (0.470)
Female (Difference)	0.0717 (0.0551)	0.350 (0.267)	0.322* (0.154)	0.380 (0.237)
Age (Mean)	−0.00111 (0.00444)	−0.00312 (0.0256)	−0.0129 (0.0156)	−0.0154 (0.0192)
Age (Difference)	−0.00352 (0.00303)	−0.0162 (0.0171)	−0.0149 (0.00894)	−0.0232* (0.0116)
Education (Mean)	0.00689 (0.0717)	−0.000663 (0.391)	−0.0581 (0.223)	−0.0843 (0.318)
Education (Difference)	−0.0378 (0.0437)	−0.206 (0.241)	−0.192 (0.141)	−0.303 (0.202)
Years of experience (Mean)	−0.0132 (0.00889)	−0.0681 (0.0490)	−0.0260 (0.0223)	−0.0379 (0.0297)
Years of experience (Difference)	−0.0000769 (0.00404)	−0.00286 (0.0222)	0.00589 (0.0120)	0.00723 (0.0162)
Puzzles free time (Mean)	0.0375 (0.0349)	0.190 (0.183)	0.105 (0.104)	0.140 (0.140)
Puzzles free time (Difference)	−0.0227 (0.0242)	−0.106 (0.120)	0.00206 (0.0919)	−0.0201 (0.127)
Understood game (Mean)	0.151** (0.0515)	0.801** (0.282)	0.436* (0.198)	0.654* (0.270)
Understood game (Difference)	−0.0273 (0.0264)	−0.160 (0.145)	−0.0637 (0.104)	−0.104 (0.139)
Knows partner (Mean)	−0.0780 (0.0719)	−0.399 (0.373)	−0.113 (0.269)	−0.199 (0.360)
Knows partner (Difference)	−0.00104 (0.0656)	−0.00334 (0.338)	−0.158 (0.197)	−0.182 (0.294)
Model type?	ols	melogit	ols	meologit
Round FE?	Y	Y	Y	Y
Dyad-type FE?	Y	Y	Y	Y
Observations	408	407	408	408

Note: Standard errors in parentheses.

* $p < 0.05$; ** $p < 0.01$.

APPENDIX H: HETEROGENEOUS TREATMENT EFFECTS BY DATE OF HIRE

	(1)	(2)
Decentralized	1.238*	0.396
	(0.591)	(0.374)
Mixed staff	1.005*	1.085*
	(0.442)	(0.474)
Old staff	0.870	0.849
	(0.659)	(0.601)
Reform # mixed	−1.459**	−1.184**
	(0.485)	(0.391)
Reform # old	−1.531 ⁺	−1.075 ⁺
	(0.750)	(0.586)
Round FE?	Y	Y
Dyad-Type FE?		Y
All Dyad controls?		Y
Constant	1.742**	0.544
	(0.532)	(2.364)
R ²	0.078	0.197
Observations	429	408

Note: Standard errors in parentheses. ⁺ $p < 0.1$; * $p < 0.05$; ** $p < 0.01$.

APPENDIX I: DRIVERS OF SUCCESSFUL BUREAUCRATIC COORDINATION (LINEAR REGRESSION)

	(1)	(2)	(3)	(4)	(5)	(6)
Decentralized	−0.592***			−0.605**	−0.515**	−0.531**
	(0.155)			(0.174)	(0.175)	(0.190)
Leadership		0.747***		0.754***		0.735***
		(0.162)		(0.157)		(0.156)
Disapproving			−1.011***		−0.935***	−0.890***
Language			(0.237)		(0.236)	(0.235)
Round FE?	Y	Y	Y	Y	Y	Y
Dyad-Type FE?	Y	Y	Y	Y	Y	Y
All Team Controls?	Y	Y	Y	Y	Y	Y
Constant	1.543	1.033	0.952	1.846	1.628	1.919
	(2.163)	(2.120)	(2.036)	(2.216)	(2.139)	(2.190)
R ²	0.189	0.214	0.199	0.231	0.211	0.251
Observations	408	408	408	408	408	408

Note: Standard errors in parentheses.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

APPENDIX J: EFFECT OF REFORM ON LEADERSHIP AND DISAPPROVING LANGUAGE

	(1) Leadership	(2) Leadership	(3) Disapproving language	(4) Disapproving language
Decentralized	0.0785	0.0167	0.0729*	0.0818 ⁺
	(0.0735)	(0.0648)	(0.0353)	(0.0465)
All team controls?		Y		Y
Dyad-type FE?		Y		Y
Round FE?	Y	Y	Y	Y
R ²	0.009	0.102	0.024	0.056
Observations	429	408	429	408

Note: Standard errors in parentheses.

⁺*p* < 0.1; **p* < 0.05.