Ideas Generation and Innovation in Bureaucracies: Evidence from a Field Experiment and Qualitative Data*

Margherita Fornasari Imran Rasul Daniel Rogger Martin J.Williams

March 2025

Abstract

We study the process of sparking innovation through ideas generation in bureaucracies, combining qualitative and quantitative evidence on workplace cultures, workplace climate, ideas generation and sharing, and bureaucratic performance. We study these issues at-scale in a developing country, working with bureaucrats in all ministries served by the Ghanaian Civil Service. Our qualitative evidence shows these organizations have strong hierarchical cultures, where juniors feel unable to raise ideas, they are not listened to when they do, or even fear being sanctioned for doing so. Our quantitative evidence comes from a field experiment training bureaucrats how to break down problems into simple solutions and raise these new ideas with colleagues. We implemented training at the individual level, and at the division-level to bureaucrats working together day-to-day. Our key finding is that individual level trainings were more effective in shifting workplace climates towards being open to new ideas, measured 6-18 months post-training. This led individuals to be more likely to raise and discuss new ideas, ultimately improving administrative processes and public service delivery. Division-level training was less effective because action plans drawn up by divisions after training failed to integrate in key aims of the intervention in terms of the nature of innovations proposed and collective steps to implementation. Rather, division-level plans reflected pre-existing workplace cultures that stifle bottom-up incremental innovations and instead, unrealistically aim for resource intensive systems-wide change. *JEL: H11,O31*.

*We gratefully acknowledge funding from the ESRC (ES/H021221/1), IGC, EDI Initiative, Stone Centre@UCL, and World Bank i2i. We thank Jane Adjabeng, Mohammed Abubakari, Julius Adu-Ntim, Temilola Akinrinade, Michel Azulai, Sandra Boatemaa, Eugene Ekyem, Paula Fiorini, Jacob Hagan-Mensah, Allan Kasapa, Kpadam Opuni, Owura Simprii-Duncan, and Liah Yecalo-Tecle for excellent research assistance, and [past] Head of Civil Service Nana Agyekum-Dwamena, Mrs. Dora Dei-Tumi, Mr. Godwin Brocke, members of our project steering committee, Andrew Wyatt, and Stefan Dercon for invaluable guidance. Special thanks to over 80 civil servants who led trainings, conducted interviews, collected data and shared insights. We thank Oriana Bandiera, Anne Brockmeyer, James Fenske, Siddharth George, Robert Gibbons, Iftikhar Hussain, Martin Mattson, John Sutton, Gabriel Ulyssea and numerous seminar participants for valuable comments. We received ethics approval from the Blavatnik School of Government Ethics Committee (SSD-CUREC1A-BSG-C1A-17-001). The trial is registered (AEARCTR-0001889). All errors are our own. Fornasari: World Bank [mfornasari@worldbank.org]; Rasul: UCL and IFS [i.rasul@ucl.ac.uk]; Rogger: World Bank [drogger@worldbank.org]; Williams: Michigan [martinjw@umich.edu].

1 Introduction

At the heart of economic progress lies the formulation and adoption of new ideas. Despite its centrality for economic prosperity, the process by which innovation is ignited through the spark of ideas generation remains poorly understood. In the macro literature, the rate of technological progress through ideas generation is either exogenously given, or endogenously determined according to some production function and the organization of R&D activities, but how ideas arise and are adopted remains unspecified [Aghion and Howitt 1998, Jones 2019]. The micro literature has instead emphasized how the arrival of ideas responds to financial rewards [Charness and Grieco 2018, Gibbs *et al.* 2019, Charness and Grieco 2023], or how ideas generation relates to individual autonomy in team contexts [Boss *et al.* 2023].[1]

We combine qualitative and quantitative evidence to study the issue in the context of bureaucratic organizations. Our qualitative evidence documents the hierarchical culture of bureaucratic workplaces, specifically with regards to their openness to ideas generation and innovation. Quantitatively, we present evidence from a field experiment that provides bureaucrats training on how to develop ideas and raise them with colleagues. We measure whether this impacts workplace climates towards fostering new ideas, actually leads to the generation of new ideas and their sharing by bureaucrats with colleagues, and ultimately whether the generated ideas are good because they improve public service delivery.

Our work builds on the literature recognizing state capacity as fundamental to prosperity [Acemoglu 2005, Besley and Persson 2010]. Much of this has examined the state through the lens of personnel economics and how bureaucratic performance is driven by the recruitment, monitoring, management, autonomy or compensation of public sector personnel [Finan *et al.* 2017, Besley *et al.* 2021]. However, scaling and sustaining high-powered incentives can be difficult in large bureaucracies [Muralidharan and Singh 2023, Williams 2025]. We build on such concerns to complement existing personnel approaches by using the lens of organizational economics, linking back to classic accounts of bureaucracies emphasizing workplace cultures as key features explaining resistance to change and low performance [Wilson 1989]. We document that the hierarchical nature of these organizations is key to stifling innovation and ideas from the bottom-up.

Our study context is a developing country in Sub-Saharan Africa: Ghana. In 2022, Ghana ranked close to the median of all countries on the World Bank's *Governance Effectiveness* scale, with its position being stable over time. It thus has many of the hallmarks of bureaucracies around the world in day-to-day challenges faced in delivering public services. We collaborated with the

[1] Central to many growth models is an ideas production function, $\frac{\dot{A}_t}{A_t} = \alpha S_t$, where total factor productivity growth – the flow of new ideas in the economy depends on research input, S_t, and research productivity, α. Research productivity is typically either exogenous or new ideas arrive at some exogenous rate [Romer 1990, Aghion and Howitt 1992, 1998]. Bloom *et al.* [2020] document that research productivity is falling over time across sectors in the US, consistent with ideas becoming harder to find. The recent literature on innovation has extended the study of the industrial organization of R&D in the direction of the micro literature, to consider the role of cultural values related to individualism, and rewards to effort and talent [Phelps 2013, Akcigit and Van Reenen 2023].

Office of the Head of Civil Service (OHCS), to understand the process of ideas generation and public service delivery at-scale: working with all central government ministries served by the Civil Service of the Ghanaian central Government.[2]

It is often stressed that bureaucracies are the classic organizational structure in which there are fundamental constraints on the provision of incentives [Dixit 2002, Besley and Ghatak 2005]. As a result, individual and collective behaviors are not entirely determined by formal rules and regulated *ex ante*, but they can also be shaped by organizational 'culture' – the classic definition of which is given by Schein [1985] as, '*a pattern of basic assumptions – invented, discovered or developed by a given group as it learns to cope with its problems of external adaptation and internal integration – that has worked well enough to be considered...the correct way to perceive, think, and feel in relation to those problems.*' At the same time, many workplaces – in public and private sectors – are characterized by hierarchical norms, where relationships between seniors and juniors also play a key role in shaping behaviors. In bureaucracies, alongside contractual inefficiencies, these hierarchical cultures shape individuals' willingness necessary to achieve consummate rather than perfunctory levels of service, and determine which of multiple equilibria arise [O'Reilly 1989, Kreps 1990, Baron and Kreps 2013]. In private sector settings, workplace culture has been shown to correlate to firm performance [Hermalin 2001, Guiso *et al.* 2015, Gartenberg *et al.* 2019]. We extend such studies to a bureaucratic setting.

Given culture is informal and can be defined in multiple ways, we use qualitative methods to measure workplace cultures and highlight hierarchical norms. In contrast, we measure workplace climate quantitatively, modifying survey instruments from settings typically focused on measuring climates of safety in healthcare [Sexton *et al.* 2006, Weaver *et al.* 2013, Martinez *et al.* 2015].[3]

Our qualitative evidence paints a stark picture of workplace cultures in Ghana's Civil Service. The core themes that emerge are: (i) there are strong cultures of hierarchy, whereby junior bureaucrats feel they are unable to raise ideas to seniors, they are not listened to when they do, or even fear being sanctioned for doing so; (ii) personal relations between juniors and superiors matters for the ability to raise ideas and be listened to. The qualitative data also makes clear there are two kinds of innovation junior bureaucrats would like to propose. On the one hand, many proposed innovations are remarkably simple, including simplifying procedures to book conference rooms, installing anti-virus software, harmonizing minute taking during meetings, moving to paperless administration, or assigning tasks more efficiently. One the other hand, pre-existing cultures are ones in which proposed innovations to improve productivity often relate to ambitious desires for system-wide improvements in the operation of the Civil Service. These are often unrealistic in

[2]The Government Effectiveness measure captures perceptions of the quality of public services, quality of the civil service and its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies [Kaufmann *et al.* 2010].

[3]Many alternative definitions of workplace culture exist including the 'stock of knowledge shared by the members of the organization' [Crémer 1993]; or a 'shared understanding among organizational members, which usually comes about through shared experience' [Weber and Camerer 2003].

terms of their resource demands, but lay the blame and focus for low productivity on actors outside the ministry. These proposals for innovation reflect more group-wide consensus rather than bottom-up ideas.

It is into such workplace cultures that we implemented a field experiment to shift workplace climates towards fostering new ideas, generating new ideas and their sharing with colleagues, with the aim of ultimately improving public service delivery.

The field experiment trained bureaucrats in three novel dimensions of ideas generation relative to standard in-service training. First, it provided them skills to diagnose and solve problems constraining productivity. This was purposefully designed to enable trainees to break down complex problems into tractable smaller issues, identify root causes, and avoid overly ambitious system-wide approaches to solving local problems. Second, the new training used motivational videos and emphasized innovation through bottom-up work processes. These videos featured actual civil servants talking about productivity routines in their divisions. Individuals were chosen from high performing divisions, providing examples of how they came up with and introduced simple innovations to improve performance. Third, bureaucrats were helped to conceptualize change and apply their new skills. We did so by asking trainees to write an action plan, in which they developed ideas for small-scale work process innovations, and how to implement them step-by-step. Trainers helped bureaucrats role play how to bring up their idea with colleagues and supervisors.

To ensure interventions were scalable, they were collaboratively designed with the Office of the Head of Civil Service, and Ghana's Civil Service Training Centre.

We experimentally implemented two versions of training. In our first treatment arm, the new training module was inserted within otherwise standard classroom training that bureaucrats usually go through as part of their promotion cycles. Within this standard training cycle, bureaucrats were randomly assigned to either: (i) standard training (C1); (ii) the new training (T1).

The second treatment arm was designed in response to the civil service's historic experience that a major constraint on the effectiveness of training was that bureaucrats often found themselves in workplace cultures in which untrained colleagues maintained a business-as-usual mentality, showing little enthusiasm for change and discouraging innovative suggestions from those coming back from standard training. To overcome this challenge, our second treatment delivered division-level training, so training groups of bureaucrats who work alongside each other day-to-day. The division-level training covered the exact same elements as the new module delivered under T1. Divisions were either randomly assigned to group training (T2) or held as controls (C2).

The relatively efficacy of these approaches in shifting workplace climate, ideas generation, and public service delivery is uncertain, and depends on factors such as the design of the intervention, the nature of problems to be solved, and workplace cultures. We discuss existing evidence from across disciplines on this in more detail later, but, focusing on the interplay of our treatments and hierarchical norms: (i) T1 might be more effective if discussion is inhibited under T2 because of the presence of Directors, hierarchical norms lead to seniors using their formal authority to

undermine the implementation of innovations raised [Aghion and Tirole 1997], or group think stifles innovative ideas being raised and reinforces pre-existing approaches to raising productivity through system-wide rather than incremental change [Benabou 2013]; (ii) on the other hand, T2 might be more effective if low-innovation norms can only be shifted away from collectively rather than individually, or if the presence of Directors at division-level training increases top-down buy-in and enables juniors to generate and share ideas [Schein 1985, Guiso *et al.* 2015].

Our main result is that the individual-level trainings were far more effective than division-level training in leading to persistent changes in bureaucracies as measured 6-18 months after training. T1 shifted forward workplace climate towards new ideas, leading trained individual bureaucrats to become more likely to raise new ideas and share them with colleagues. Ultimately this fed through into improvements in administrative processes and public service delivery.

More precisely, individual training led to a significant increase in an overall index of bureaucrat's perceived workplace climate ($p = .009$). This was driven by index subcomponents related to fostering new ideas ($p = .037$), teamwork ($p = .083$) and performance ($p = .023$). These all link to the new training given its emphasis on innovation, teamwork and improving work processes. This shift in climate translated into the generation of new ideas by bureaucrats and ideas sharing. Relative to bureaucrats assigned to the standard classroom training (C1), those assigned to the new training were significantly more likely to report: (i) freedom to express thoughts, feelings, and criticisms ($p = .030$); (ii) being able to raise suggestions around procedures ($p = .046$); (iii) being able to discuss processes to improve workplace productivity ($p = .029$).

Finally, to understand whether the ideas generated are 'good' ideas, we link our treatments to measures of bureaucratic performance. To measure performance related to core activities such as planning, oversight and policy formulation, we undertook two exercises in coding administrative data. These relate to: (i) the quality of administrative processes; (ii) task completion by division. We find that administrative files from divisions with at least one member assigned to T1 had significantly higher quality procedures ($p = .054$) and adherence to procedures ($p = .050$). On public service delivery, we find that task completion rates rose for divisions that have at least one individual assigned to T1 ($p = .071$): the magnitude of the effect corresponds to a 68% increase in tasks being fully completed.

In contrast, comparing bureaucrats in divisions randomly assigned to division-level training (T2) to those in control divisions (C2) we find null (or sometimes negative) impacts on nearly all margins of workplace climate, ideas generation, administrative processes and task completion.

We then investigate why individual and division-level training had such contrasting impacts. We do so by first analyzing action plans drafted at the end of each intervention. While the drafting of action plans was standard practice and continued in C1 trainings, action plans in T1 and T2 were re-designed to help bureaucrats better conceptualize change and actually apply their new skills and motivations in the workplace. The plans were produced by trainees on a topic of their choosing, but with guidance, coaching, and feedback from trainers. Each individual in

the standard and individual new training developed their own action plan. In the division-level training, the division developed a single action plan,

We find that division-level training was less effective because the action plans drawn up by divisions collectively failed to integrate in the key ideas of the novel training. As a result, they focused on non-core work processes, suggested fewer follow-up meetings to implement ideas, and were harder to achieve in the sense that they required acquiring additional resources, the provision of further training, and a greater need to meet with actors outside the organization. In this sense, the division-level action plans in T2 much more reflect pre-existing workplace cultures, where training was often seen to be ineffective because constraints on productivity were due to service-wide inefficiencies, and thus a lack of resources was cited being the single largest obstacle to addressing performance issues. This is very far from the spirit of the training provided (and internalized under T1) where individuals/divisions were encouraged to break down complex problems into manageable solutions.

We find far weaker evidence for alternative mechanisms for the ineffectiveness of T2 such as: (i) the presence of Directors at the division-level training inhibiting the raising of innovative ideas or slowing down the adoption of new work processes; (ii) heterogeneity within divisions making it harder to drive forward changes in workplace climate and an increased openness to innovation and new ideas.

Taken together, our results echo findings that the impact of interventions in organizations can depend critically on pre-existing workplace cultures – where we highlight novel insights arising from the fact that organizations can be hierarchical [Blader *et al.* 2020, Celhay *et al.* 2024].

Our key contributions of general interest are twofold.

First, our analysis breaks new ground in using a field experiment at-scale to shift workplace climates in bureaucracies. We complement existing work on the personnel economics of the state [Finan *et al.* 2017, Besley *et al.* 2021], and bridge to the growing literature on non-monetary incentives in private sector organizations, such as social incentives [Bandiera *et al.* 2009, Ashraf and Bandiera 2018], relational contracts [Blader *et al.* 2015, Adhvaryu *et al.* 2024], climates against harassment [Alan *et al.* 2023], and worker voice [Heining *et al.* 2021]. We link to these developing strands of work as we present our findings.

Second, we provide insights on the mechanics of innovation and ideas generation in organizations. We build on correlational evidence from other disciplines suggesting organizational culture drives innovation and performance [Ahmed 1998, Jassawalla and Sashittal 2002], but organizational dynamics can also inhibit the generation and adoption of ideas [Amabile 1996, Edmondson 1999, Ashford *et al.* 2009, Fast *et al.* 2014]. We bring these ideas to the field in a real world bureaucratic setting, where our focus on low- and mid-level bureaucrats advances the study of bottom-up innovation and voice in bureaucracies [Fernandez and Moldogaziev 2012, Hassan 2015]. Our findings hinge on two key features of our context: contractual inefficiencies abound [Dixit 2002, Besley and Ghatak 2005], and organizations are hierarchical, both of which leave scope for workplace cul-

ture to be a key driver of individual and collective behavior [O'Reilly 1989, Kreps 1990, Baron and Kreps 2013]. All organizations – in public and private sectors – are subject to both features to some extent, so this provides a four-way classification for future work to explore in more detail.

The paper is organized as follows. Section 2 describes our study context, and the quantitative and qualitative data collection underlying our analysis. Section 3 describes the individual and division-level trainings, and the experimental design. Section 4 presents our core results on how each treatment impacted bureaucrats' perceptions of workplace climate, ideas generation and sharing, the quality of administrative processes and bureaucratic performance. Section 5 explain why division-level training was relatively ineffective. Section 6 concludes by discussing how our findings open up a future research agenda. Appendix A presents additional results, Appendix B details supporting extracts from qualitative interviews and Appendix C shows key survey modules.

2 Context and Data Sources

2.1 The Ghanaian Civil Service

The Civil Service of Ghana's central government comprises 45 ministries. They perform core bureaucratic functions related to policy making, administration, and overseeing public service delivery. Each organization has defined responsibility for a sector and tasks. Tasks to be undertaken in a given year are determined via an annual planning, budgeting, and reporting cycle, in which long-term strategic plans and government priorities are translated into work programs. Organizations are staffed almost exclusively by career civil servants, with the exception of politically appointed ministers and a handful of special advisors. Recruitment, staff assignment, transfers, promotions, and termination are handled centrally by the Office of the Head of Civil Service (OHCS). As a result, there is *de jure* uniformity of human resource policies across organizations.[4]

Within each organization responsibility for tasks is distributed among divisions (teams). Organizations typically have between four and ten divisions. Four divisions form the spine of every organization: Finance and Administration; Policy, Planning, Monitoring and Evaluation; Research, Statistics, and Information Management; and Human Resources. Other divisions are organization-specific. Each division comprises 6 to 15 civil servants and is headed by a Director.

Table A1 details the 45 government ministries, ranked by their annual budget. 38 ministries form our core sample, including 26 of the 27 largest budget ministries. They cover all the core functions of state and have an aggregate budget of over $5bn.

[4]Our study sample covers all ministries and departments forming Ghana's Civil Service (as opposed to the broader Public Service), with the exception of the Office of the President. The Service excludes frontline public service delivery agencies, such as the Ghana Education Service, as well as local government employees and staff of semi-autonomous agencies, such as the Bank of Ghana.

2.2 Data Sources

We use multiple data sources to construct a detailed timeline of training delivery, and its impacts on workplace climate, ideas generation and public service delivery. Figure 1 shows the study timeline, indicating how the key stages of data collection relate to the timing of trainings.

2.2.1 Bureaucrat Surveys

A key source of data are surveys we fielded to individual bureaucrats. We implemented bureaucrat surveys at baseline from August to November 2015, and at endline from June to October in 2018. The endline takes place 6-18 months after training. Given the specialized and sometimes sensitive nature of questions, enumerator teams comprised former civil servants. We ensured no one interviewed those personally known to them, and held enumerator teams largely constant between baseline and endline.[5]

Our baseline survey covered 2971 bureaucrats (corresponding to 92% coverage of professional staff lists), and our endline survey covered 3297 bureaucrats (84% coverage).[6]

2.2.2 Workplace Climate

One survey module measured perceived workplace climate. To do so, we adapted the scales used by Sexton *et al.* [2006] and Martinez *et al.* [2015]. These scales were developed in contexts of healthcare and aviation – both settings where hierarchical norms can worsen performance. The scales have established sub-indices related to teamwork climate, performance climate, stress recognition, perceptions of management, and working conditions. We construct (modified) versions of these sub-indices and also measure a new dimension of climate: fostering new ideas. Table 1 shows the questions asked under each sub-index at baseline and endline, all of which respondents were asked to report on a five-point Likert scale. Table C1 makes precise the comparison between our questions those originally used by Sexton *et al.* [2006].[7]

[5] A survey team comprising a research assistant and five enumerators visited each organization for a continuous period of anything from several days to several weeks (depending on its size), conducting face-to-face individual interviews in private rooms. Enumerators explained the study goals, that participation was voluntary, and that individually identifiable data would be available only to the research team and not to the Civil Service. Survey teams made extensive efforts to interview every individual on the staff list, even if they were unavailable or were called to attend to work duties on short notice on the day initially planned for interview.

[6] Civil servants are formally divided into senior and junior staff, with the distinction depending on educational qualifications and career track. Senior staff comprise the core of officials in professional and administrative positions, usually with a university degree or post-graduate qualification, while junior staff comprise support staff like drivers, secretaries, and cleaners. The staff lists that we used focused on senior staff. Our survey covers the full spectrum of rank and tenure within the senior service, from recent entry-level hires to division Directors.

[7] Sexton *et al.* [2006] developed their scale in the context of hospital care and defined safety climates as the 'product of individual and group values, attitudes, perceptions, competencies and patterns of behavior that determine the commitment to, and the style and proficiency of, an organizations health and safety management'. This was itself based on an earlier measurement of organizational climate used in aviation safety. Weaver *et al.* [2013] provide a meta-analysis of interventions aiming to improve safety climate in healthcare: this shows the scale to be validated in that changes are correlated to improvements in patient outcomes.

The new training module relates closely to sub-indices on: (i) fostering new ideas – where we questioned bureaucrats about their suggestions on workplace productivity being acted upon by management, the speed of adoption of new ways of working in the division, and their ability to see ways of improving work in the division; (ii) teamwork climate – where we asked respondents about their ability to ask questions, resolve disagreements, coordinate between juniors and seniors, and the frequency of communication breakdowns; (iii) performance climate – where we asked about workplace culture making it easy to learn from others, knowing the proper channels through which to direct questions about bureaucratic processes, the provision of feedback, the handling of errors, being encouraged to raise concerns, and feeling happy to be a citizen served by the division.

The raw values of each sub-index at baseline are shown in each row of Table 1. Each sub-index – including that for fostering new ideas – shows more variation in climate within divisions in the same organization than between organizations.

We aggregate these into an overall index of workplace climate by taking the division mean of each variable, z-scoring each variable using the cross-division means and standard deviations, and computing aggregate indices by averaging z-scores. Throughout, the aggregate organizational climate index is presented as a normalized z-score. Figure 2A shows how aggregate workplace climate varies across and within organizations (i.e. across divisions in the same organization). We rank organizations in increasing level of the overall climate index.

We observe considerable variable in the workplace climate of organizations and divisions within them. This is despite the fact that there is *de jure* uniformity of human resource policies across organizations, and they recruit from the same talent pool. It is also despite these organizations being in close geographic proximity and having opportunities to share knowledge. This opens up a role for informal channels – such as workplace cultures and relational contracts – to help explain what shapes workplace climate and attitudes towards innovation in otherwise similar organizations and divisions [Gibbons and Henderson 2012, Besley and Persson 2024].

2.2.3 Qualitative Data

To provide insight into how workplace cultures relate to innovation, we conducted qualitative interviews in parallel to our quantitative data collection. Interviews were conducted in a semi-structured fashion, so guided by a consistent set of questions but the amount of time spent on each topic varied depending on interviewee responses, and interviewers could flexibly probe and follow-up. Interviews were conducted anonymously and were not audio recorded to ensure interviewees could speak openly and frankly, with handwritten notes that were subsequently typed up. Quotes from interviewees presented are thus paraphrases rather than verbatim transcripts.

We conducted 51 interviews with bureaucrats from nine divisions in ministries. Divisions were selected at random from the set of Policy, Planning, Monitoring, and Evaluation divisions – one of the four common divisions in all ministries. To assess how workplace cultures were impacted

by the treatments, divisions in the qualitative data collection exercise were stratified by treatment status. This process yielded six divisions whose members had some combination of trainings, and we deliberately over-sampled a further three divisions from those that had no members undergo standard training during the year to help understand innovation dynamics in the absence of the interventions. For each of these nine divisions, we interviewed every member of the division from the Director to the least senior member. Interviews were conducted between November 2017 and January 2018, towards the end of the training roll-out. We conducted 22 follow-up interviews in January 2018 to further probe themes that had emerged.

The topics covered included workplace relations, productivity, and ideas generation in the division. The initial set of questions prompted general feedback from respondents, unrelated to training to avoid priming and social desirability bias, and to draw out long-standing views about workplace cultures and how they impacted innovation. Later parts of interviews were designed to elicit qualitative feedback specifically related to the treatments, actions plans developed and perceptions of how trainings undertaken (if any) had affected them at the individual and division-level in terms of driving forward innovation.

2.2.4 Bureaucratic Performance

To understand whether the ideas generated are 'good' ideas, we link our treatments to measures of bureaucratic performance. A challenge is that the majority of tasks conducted by bureaucrats are not related to tangible outputs – procurement and physical project construction are exceptions, not the rule [Rasul *et al.* 2021]. To measure performance related to core activities such as planning, oversight and policy formulation, we undertook two extensive exercises in collecting and coding administrative data. These relate to: (i) the quality of administrative processes; (ii) task completion by division. These data sources allows us to create actual performance measures separate from our surveys, and hence that are not subject to concerns over social desirability bias in bureaucrat reporting of public service delivery.

Quality of Administrative Processes Filing and record-keeping are basic but crucial tasks of bureaucrats. To measure these processes and how they were impacted by our treatments, we collaborated with OHCS and the Public Records and Archives Administration Department (PRAAD), utilizing the following procedure. First, our enumerators visited division offices unannounced, requesting to see file registries. These registries list all files that have been acted upon by division members including: (i) open files dealing with currently active projects; (ii) closed files related to completed projects. Our enumerators selected two open and two closed files from the registry list, drawing evenly from the study period (i.e. files opened from 2015-18) and excluding files that had closed pre-intervention. We selected files from the four divisions that form the spine of every organization – Finance and Administration; Policy, Planning, Monitoring and Evaluation;

Research and Statistics, and Human Resources. Our sample covers 286 files from 106 division in 37 organizations.

We then worked with a team of retired senior civil servants to code the quality and completeness of each file, according to a scoring grid designed in conjunction with expert civil servants (Table C2). We create division-level measures of administrative process along two dimensions: (i) the quality of procedures: capturing whether files complied with rules about how they should be handled, compiled, and circulated (Panel A of Table C2); (ii) the quality of content: capturing whether the information compiled, analysis undertaken, and decisions made within files reflect good administrative practice as expected by the Civil Service (Panel B of Table C2).

Task Completion by Divisions Organizations are required to submit quarterly and annual reports detailing completion of tasks from their annual workplan. These reports span the spectrum of bureaucratic activity from procurement, personnel management, infrastructure development, policy development and analysis, permitting and regulation, and financial management. Our research team worked with civil servants from OHCS to compile these reports, extract and standardize data and code the completion of each task on a continuous scale from one to five, with a score of one representing 'No action was taken towards achieving the target', a score of three representing a task for which 'Some substantive progress was made towards achieving the target – the task is partially complete and/or important intermediate steps have been completed', and a score of five corresponding to 'The target for the task has been reached or surpassed.' We code the completion of 1473 tasks at the division-level across 38 organizations for the endline year 2018, so 6-18 months after training is completed. We thus establish medium run impacts of the treatments on bureaucratic performance.[8]

To validate our task completion data and measure of organizational climate, we used our baseline survey data to examine the relationship between the two. Figure 2B shows a strong positive association between task completion rates and organizational climate. While the observational nature of the data means that this positive relationship is not necessarily causal, it is nonetheless reassuring that our adapted organizational climate index captures features relevant for public sector performance in our study context. We establish causal impacts using our field experiment.[9]

[8] OHCS prescribes a template containing a list of tasks, projects, outputs, and processes the organization has planned to undertake during the specified time period, together with actions taken during the reporting period.

[9] There might of course be other factors that drive variation in performance. As discussed in Blader *et al.* [2015], these include: (i) perceptions – organizations are blind to innovation; (ii) inspiration – organizations do not know what to do about performance gaps because practices underlying performance involve tacit knowledge and are hard to learn and communicate; (iii) motivation – organizations do not care about poor performance because they lack competitive pressures to improve.

2.3 Data Linkage

Linking individuals across data sources is challenging because unique identifiers for bureaucrats in the Civil Service were not in common use during our study period, and there can be multiple spellings of names. This creates serious challenges in linking individuals from baseline to endline so we focus on tracking individuals from training to the endline bureaucrat survey. Table A2 shows correlates of attrition: Columns 1 to 3 focus on attrition between T1 and C1, and Columns 4 to 6 examine attrition of those in T2 relative to those in C2. Of the individuals that attended C1/T1 trainings, we track 86% of them to the endline bureaucrat survey. Of the individuals that were in the 152 T2 divisions, we track 88% of them to endline. For each treatment we examine whether attrition is: (i) unconditionally correlated to treatment; (ii) is correlated to treatment conditional on bureaucrat characteristics; (iii) heterogeneous across individuals depending on these characteristics. For both T1 and T2 we see that attrition is uncorrelated to treatment, and nor do we find evidence of heterogeneous attrition – the p-values on a test of the joint significance of all interactions with treatment T1 and T2 are .335 and .902 respectively.

Divisions can also be difficult to link over time because of changes in division titles. Of the of the 152 T2 divisions, we are able to link to administrative process data for 106 (70%) of them, and link to endline assessments of bureaucrat performance for 84 (55%) of them.

2.4 Workplace Cultures

We use our qualitative evidence to establish key elements of bureaucratic culture related to ideas generation and innovation.

2.4.1 Hierarchy

Table B1 presents sample extracts from the qualitative interviews in response to questions related to bureaucrats generating ideas to improve workplace productivity. We draw out three broad themes, limiting ourselves to five extracts per theme.

The first theme is the strong culture of hierarchy, whereby junior bureaucrats often feel they are unable to raise ideas to seniors, they are not listened to when they do raise new ideas, or even fear being sanctioned for raising ideas [Williams and Yecalo-Tecle 2020]. As stated by an interviewee, '*After expressing [an idea] they may say 'I'll call you later' and that just ends the idea there... So why force myself to try if it will be ignored? I usually get 'we will talk later' and when I go, [superior's name] postpones the time saying 'tomorrow'...[superior's name] then tells me that I am putting pressure... In the civil service before you take any action the powers can jump on you and you suffer that alone. Sometimes you go through transfers or they query you by giving you a note on why you are doing this and to explain. Things can then go on your file.*' (A2)

The second theme is the importance of personal relations between juniors and superiors in

being able to raise ideas. As stated by an interviewee, '*I hear from colleagues that ideas are rejected and this lowers your morale. It all depends on if it is rejected in a nice way or is it done outright... This all depends on the superior you have and it boils down to the relationship you have with your superior. If you have a superior with a listening ear it is always good to think through things before you share it... This also all depends on the superior's character.*' (A11)

The third theme builds on this, with extracts describing how ideas spread from the ground up: '*You need to do your homework first and get the support from officers. The initial conversations about an idea start on the ground and then it is raised up... Having the support with colleagues really helps.*' (F2); '*One aspect is the organizational culture... If you go to an institution where they do not listen to newly recruited staff, they do not have a voice. The civil service does not help in that manner as they make you feel that young people are not knowledgeable. It is not said, but in discussions, in the aura and ambience, you feel it. They may just say 'this is the way this is done' but you may realize that this is not right. I have personally suffered with this issue and I know other juniors who have suffered it.*' (A6)

However, we do not want to paint a uniformly negative picture of ideas being stifled. To reflect this, we also draw out are qualitative extracts that reflect positive experiences of being able to raise and implement new ideas. For example one interviewee stated, '*When there is a new idea, we discuss it and take it up. Seniors do encourage us during meetings. There is a channel in which you go through the director to the Chief Director.*' (D5).

2.4.2 Examples of Innovation

To make concrete the kinds of bottom-up innovations that bureaucrats typically propose, Table B2 presents sample extracts related to ideas raised. In contrast to common negative stereotypes of bureaucrats, officials do have meaningful ideas for improving performance. We draw out two broad types of ideas raised. Row A lists examples of the remarkably simple proposed innovations: these include simplifying procedures to book conference rooms, installing anti-virus software, running informal trainings during lunch times, or harmonizing minute taking during meetings. In contrast, Row B presents ideas relating to more ambitious system-wide changes in operations, typically involving managing engagement with multiple stakeholders outside the ministry.

The kinds of simple innovations proposed in Row A dovetail more with what is actually emphasized in the training interventions, where individuals/divisions are encouraged to break down complex problems into the adoption of manageable solutions.

Implications The extracts on workplace culture and ideas generation have implications for the efficacy of the training interventions. The kinds of pre-existing workplace culture that exist suggest there can be variation in treatment effects of division-level training depending on precisely how hierarchy affects ideas generation. If hierarchical norms lead to a strong pre-existing consensus

around what kinds of action are needed to overcome productivity constraints and bottlenecks, that can lead divisions to fall back into group-thinking and suggest more resource intensive system-wide changes in operations. On the other hand, if hierarchical norms are enforced through seniors, then the presence or absence of Directors at division-level training can have implications for the kinds of ideas that are generated, although hierarchical norms might then lead to such ideas not being followed up and implemented. We tease apart these channels when examining mechanisms.

3 Treatments

It is into such workplace cultures that we inject novel trainings, with the aim of shifting workplace climate and processes of ideas generation, to ultimately improve public service delivery. We first describe the business-as-usual form of classroom training for bureaucrats, and how we build on this structure with these our treatments.

3.1 Training of Bureaucrats

Standard Classroom Training It is standard practice for civil servants to take part in training courses as part of their promotion cycle. These cycles occur every few years, so that in any given year, around 15% of bureaucrats are due to attend standard classroom training. These training courses last ten days and are tailored by civil servant grade. Given the link to promotion, individuals have strong incentives to attend and individuals typically return back to the same division post training, having moved up one grade.

The training is developed and delivered by in-house government trainers – the Civil Service Training Centre (CSTC). CSTC publishes an annual calendar of trainings. There were 17 10-day training sessions scheduled during February-December 2017, our period of intervention roll-out. Organizations decide which session to send individuals to. The average classroom size is 20, with co-attendees being bureaucrats of the same grade but from different organizations and divisions.

Table C3 lists the modules covered in the standard training: on administrative writing, policy development and analysis, administrative principles and instructions, human relations, organizational security and safety, team building, work ethics and work standards, and productivity. The standard productivity module focuses on defining and measuring productivity, and adopts a Japanese *kaizen* approach to continuous improvement. It is fair to say that the module is widely perceived to be outdated, abstract, and providing few context-specific examples on how to practically improve productivity.[10]

[10]The *Kaizen* approach to creating continuous improvement was originally developed in the context of Japanese manufacturing firms. This approach to improving work processes lay at the heart of the intervention evaluated at-scale in Muralidharan and Singh [2023] that was designed to improve management quality in Indian schools – but was demonstrated to have had no measurable impact on school functioning or student outcomes.

A New Productivity Module In the year prior to our intervention roll out, we collaborated with OHCS and CSTC to design a new, more interactive and useful module on day-to-day workplace productivity, to insert into the otherwise standard classroom training courses. This was developed using evidence from our 2015 baseline survey of bureaucrats that documented: (i) organizations delegating their staff with greater autonomy, discretion, and flexibility are associated with higher rates of task completion; (ii) creativity and tacit knowledge of lower- and middle-level bureaucrats is important for productivity [Rasul *et al.* 2021]. While we provided input on the new module, the curriculum was drafted by an international development consultant. OHCS and CSTC ensured it was well suited to the context.

The new module had three novel and interwoven components.

First, it provided skills to diagnose and solve problems constraining productivity. We presented evidence from our baseline survey that some organizations were more productive than others, discussed reasons for poor productivity, and prompted trainees to think of examples of productivity measurement and bottlenecks in their own and other organizations. It then introduced four problem-solving skills: problem-tree analysis, force-field analysis, fishbone diagrams, and the 'five whys'. This allowed trainees to break down complex problems into smaller tractable issues, identify root causes, and avoid overly ambitious system-wide approaches to problem solving [Andrews *et al.* 2017]. Trainers helped trainees apply these methods in their own work context.

Second, the new module used motivational videos and emphasized innovation through bottom-up work processes. These videos featured real life civil servants talking about productivity routines in their division. Individuals were chosen from high performing divisions, providing examples of how they came up with and introduced incremental innovations to improve performance. A segment of the videos also included the Head of Civil Service and senior officials emphasizing the importance of innovation and productivity in the bureaucracy.[11]

The third component helped bureaucrats conceptualize change and actually apply their new skills and motivations in the workplace. We did so by asking individuals to develop an action plan, in which trainees developed ideas for small-scale work process innovations, and how to implement them step-by-step in their division. To give voice to bureaucrats, trainers helped bureaucrats role play how to bring up their idea with colleagues and supervisors.

Trainees were prompted to focus on problems that were small and specific enough to be actionable within their division, rather than on structural problems or on actions requiring significant resources or approval from higher authorities – so aiming to shift away from typical workplace cultures that focused on systemic-approaches to innovation. Action plans were produced by trainees on a topic of their choosing, but with guidance, coaching, and feedback from the trainers. While

[11] Videos were seen as a far more scalable intervention than study visits or other in-person approaches across organizations. The video was 20 minutes long and produced by an Accra-based video producer with guidance from CSTC and the research team. Trainers facilitated discussion after each part of the video, asking trainees to relate the content back to their experiences, and referred to it throughout the latter part of the training.

action plans existed before our intervention, we modified their design to make them more precise, and interlink with the problem solving skills training provided in the new module. Table C4 contrasts the old and new style action plans.

Overall, the new productivity module aimed to improve both the problem solving skills and motivation of civil servants, emphasizing the importance of individual initiative (in contrast to hierarchical workplace cultures), and help participants apply incremental innovations in their day-to-day work – even down to role-playing how they would bring up their new ideas with their colleagues and superiors.

3.2 Treatments and Design

Our field experiment was designed to use the new productivity module to shift workplace cultures, promoting a climate of ideas generation and sharing. We implemented two treatments in a 2×2 factorial design.

T1: Individual Classroom Training In our first treatment arm, within the standard 10-day training course, we replaced the standard productivity module with the new productivity module. As shown on the study timeline in Figure 1, the new productivity module was rolled out during the 2017 training calendar. The final row of Table C3 shows the new productivity module inserted as part of the otherwise standard training curriculum. All other modules remain unchanged. The new module took place on last/penultimate day of courses. At the start of this day we randomly assigned individuals to either: (i) the standard productivity module (C1); (ii) the new productivity module (T1), with an equal split of bureaucrats across C1 and T1. In 11 of the 17 training sessions scheduled during February-December 2017 we were able to randomly assign bureaucrats to T1 or C1, so these sessions serve as a our randomization strata. In the other sessions bureaucrats followed the standard curriculum (C1) because all classrooms at CSTC were fully occupied – and these six sessions are spread over the training cycle. In total, 139 bureaucrats were assigned to T1 and completed the new productivity training model, while 283 completed the standard training.

The treatments were blinded: bureaucrats did not know which training they were receiving (C1 or T1). Both sets of trainings we delivered by CSTC on CSTC-branded materials. Individual trainers delivered either C1 or T1 (to avoid contamination), and enumerators monitored any sharing of materials between T1 and C1.

T2: Division Training The second treatment arm was designed in response to CSTC's historic experience that a binding constraint on training effectiveness was that bureaucrats often found themselves back in workplace climates in which non-trained colleagues maintained a business-as-usual mentality, showing little enthusiasm for change and discouraging innovation from trainees.

To overcome this challenge, our second treatment delivered division-level training. At the

end of each 10-day training course, 40% of trainees were randomly selected (from T1 and C1) and informed they and their division members would be given a day long training session. This covered the same three novel elements of the new productivity module. All grades within the division were invited to participate, including Directors. Division-level trainings were delivered by the same CSTC trainers that delivered the T1 trainings, and as shown in Figure 1, took place from April 2017-March 2018, so with a built-in lag ensuring division training occurred around three months after individual training courses were completed. We refer to the division training as T2, with the experimental counterfactual being the 60% of trainees (from T1 and C1) whose division were not selected for division training (C2). In C2 teams, no division-level meetings took place, so there was no form of placebo division interaction.[12]

Design Figure 3 shows the experimental design. Panel A focuses on classroom training. Of the 2971 bureaucrats in service in 2017, 447 (15%) were up for promotion and so required to attend CSTC training. Our sample includes 422 (94%) of them: 283 are randomly assigned to the standard curriculum (C1), and 139 are assigned to the new productivity module (T1). Our sample includes a further 2524 bureaucrats not up for promotion training, that we refer to as group C0.

Panel B details the division training. Of the 416 divisions across organizations, 152 have at least one bureaucrat up for promotion training, and so assigned to either C1 or T1. Of these, 732 bureaucrats in 93 divisions (from 34 organizations) were assigned to T2, and 540 bureaucrats in 59 divisions (from 36 organizations) were assigned to C2.

Panel C shows the training combinations: 189 bureaucrats are assigned to C1 and C2 (and of these 173 or 92% are sampled), 112 bureaucrats are assigned to C1 and T2 (98% are sampled), 83 bureaucrats are assigned to T1 and C2 (93% are sampled), and 63 bureaucrats are assigned to T1 and T2 (99% are sampled).

Treatment Comparisons Both treatments deliver the same novel productivity module, and are delivered at scale by standard in-house civil service trainers. The key differences are that: (i) in T1 individuals are trained alongside others in the same grade but not their colleagues; (ii) in T2 individuals are training alongside division colleagues; (ii) T1 takes place in the context of the standard training course, while T2 offers the new productivity module as a standalone; (iii) in T1 each individual developed their own action plan, while in T2 the entire division developed a single action plan.

Little is causally established on the relatively efficacy of these approaches in shifting workplace climate and ideas generation. The related literature from across disciplines suggests differential impacts could either way due to the design of the intervention, the nature of problems to be solved,

[12]The 40% probability was calculated by simulation pre-intervention, as the rate most likely to equalize sample sizes across treatment cells given that once a division was selected for training, it was removed from subsequent randomizations. The aim was deliver 100 T2 trainings by March 2018. We achieved 93 division training sessions.

or workplace cultures.

On design issues, T2 training might be more effective if it is tailored to actual work that divisions are tasked with, or if the production function in divisions exhibits strong complementarities in co-worker effort. Alternatively, T1 might be more effective because it allows individuals to lead their division in change, avoiding miscoordination and expectation errors across co-workers [Hermalin 1998, Bolton *et al.* 2013].

On the nature of problems, T2 might be more effective for tasks requiring complex solutions, because a group is more likely to reject incorrect solutions. On the other hand, heterogeneity within divisions might lead to disagreement about the specific ideas to take forward [Charness and Greico 2019].

The kind of workplace culture the treatments are injected can also matter – depending on how hierarchical norms impact ideas generation. T1 might be more effective if discussion is inhibited under T2 because of the presence of Directors, hierarchical norms lead to seniors using their formal authority to undermine the implementation of innovations raised [Aghion and Tirole 1997], or group think stifles innovative ideas being raised and reinforces pre-existing approaches to raising productivity through system-wide rather than incremental change [Benabou 2013].

On the other hand, T2 could be more effective if workplace cultures represent group equilibria, that can only be shifted by coordinating across group members [Chassang 2010, Gibbons and Henderson 2012, Baron and Kreps 2013], or if the presence of Directors at division-level training increases top-down buy-in and motivates juniors to generate and share ideas [Schein 1985, Rotemberg and Soloner 2000, van den Steen 2005, Guiso *et al.* 2015].

3.3 Views of Training

Qualitative Evidence Qualitative feedback on the trainings is summarized in Table B3. These interview extracts are in response to questions about the effectiveness of training at helping civil servants come up with new ideas for improving productivity. We did not prompt interviewees about the distinction between different trainings. Standard training (C1) is viewed as useful – our treatments build on this. The treatments were viewed as being innovative and providing ways to help juniors come up with relatively simple to implement ideas on productivity improvements. The motivational videos in T1 and T2 proved to be extremely popular, with many citing it as their favorite component, and referring back to it in discussions.

Quantitative Feedback Table A3 presents quantitative evidence on bureaucrats engagement with each training (C1, T1 and T2), as elicited from facilitators who observed training days. T1 and T2 are seen to have more engagement from bureaucrats than standard training C1, with the division training T2 having higher levels than for the classroom training T1. Team work is not utilized at all in the standard training, while 80% of observers note such methods being used in

T1 and T2. Role play is also not utilized in C1, while around 30% of observers note such methods being used in T1 and T2. The practicality of T1 and T2 is also seen to be higher than standard training, with a greater use of examples. T2 especially is seen to have employed task-specific language – perhaps reflecting its greater ability to be tailored to the needs of divisions. Finally, we see little evidence of differences in the quality of delivery across C1, T1 and T2.[13]

3.4 Balance

Table 2 examines balance on bureaucrat and division characteristics between those assigned to treatment and control for the classroom training: T1 vs. C1. As Panel A shows, controls are equally split by gender, their average age is 38, and the majority have an undergraduate degree. Individuals up for promotion have served in the service for a decade. We see from Column 3 that within randomization strata, samples are well balanced across individual and division characteristics. The one exception is gender, so we condition on that in our analysis related to T1/C1. Panel C shows perceptions of aggregate workplace climate, and related to fostering new ideas specifically, are balanced.

As a point of comparison, Column 4 shows characteristics of bureaucrats not up for promotion training in the 2017 cycle (so were never eligible for either C1 or T1). We denote this group as C0. Column 5 shows the non-experimental comparison between C0 and C1, highlighting that as expected, younger and less experienced bureaucrats are up for promotion.

Table 3 examines balance for the division-level intervention. Columns 1 and 2 show division characteristics for those assigned to control divisions (C2) and into division training (T2). Control divisions tend to be smaller ($p = .079$), and with a lower share of female bureaucrats ($p = .036$). We account for these imbalances where relevant in our analysis. Panel B shows some slight imbalance on the aggregate index of workplace climate. We later examine the robustness of our baseline result on how T2 impacts workplace climate when additionally controlling for baseline measures of climate.[14]

[13]The notion that the treatments were equally effectively implemented is reinforced by data collected from bureaucrats participating in T1 and T2, where we find: (i) 73% (86%) of respondents agree that training met their objectives; (ii) 82% (80%) agree that the sessions were relevant to their needs; (iii) 83% (87%) agreed that trainers delivered the training well.

[14]The division level statistics in Table 3 include all bureaucrats, including those not up for promotion in the 2017 cycle. This is why the average years in service are higher in Table 3 than in the Table 2 for those bureaucrats assigned to either C1 or T1.

4 Results

4.1 Workplace Climate

We first estimate impacts of each treatment on perceived workplace climate. This outcome is measured in our endline civil servants survey, some 6-18 months after training takes place. For individual classroom training, we consider bureaucrats assigned to either C1 or T1 that we track from training to the endline survey, and estimate the following specification for individual i:

$$work_climate_{id} = \alpha_0 + \alpha_1 T1_{id} + \beta \mathbf{X}_{id} + u_{id}. \quad (1)$$

The outcome $work_climate_{id}$ is the workplace climate index described in Table 1. In $\mathbf{X}_{id}$ we control for gender (the one dimension of imbalance) and the randomization strata. Standard errors are clustered at the division-organization level.

The results are in Panel A of Table 4, where at the foot of each Column we report the mean of the index for those in C1 tracked to endline. In Column 1 we see that the overall index of division culture rises by $.180\sigma$ ($p = .009$). The remaining Columns examine index sub-components. The teamwork climate sub-index significantly rises by $.189\sigma$ ($p = .083$), the performance climate sub-index rises by $.191\sigma$ ($p = .023$), and the fostering new ideas sub-index rises by $.194\sigma$ ($p = .037$). These longer term changes all link closely to the new training module given its emphasis on innovation, work process improvement, and teamwork. However, T1 also shifts forward other dimensions of workplace climate: perception of management increases by $.156\sigma$ ($p = .063$), and perceptions of working conditions rise by $.179\sigma$ ($p = .090$).

We next repeat the analysis for those assigned to division training, where $T2_{id} = 1$ if the division d is assigned to receive division training, and zero otherwise. The comparison is to those whose divisions were randomly assigned as controls (C2). In $\mathbf{X}_{id}$ we control for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division, whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division. We continue to cluster standard errors at the division-organization level.

The results are in Panel B of Table 4, where again at the foot of each Column we report the mean of the index for those in C2 tracked to endline. In Column 1, we see no shift in the aggregate index of workplace climate. The remaining Columns show that no sub-index of climate is shifted by the division-level training including the sub-index on fostering new ideas.

In Table A4 we present robustness checks related to sample selection. In Panel A we re-estimate the impacts of T1 for the sample of bureaucrats linked from baseline to endline, to check if changes in perceptions are driven by the composition of bureaucrats. Despite the reduced sample size, we continue to find significant increases in the fostering new ideas sub-index ($p = .015$). In Panel B, we re-estimate the impacts of T2 but redefine the treatment dummy $T2_{id}$ to be one if bureaucrat

i attended the division level training T2 (rather than their division being assigned to treatment). We continue to find null impacts throughout. In Panel C we repeat our baseline estimate of the impact of T2 but given the baseline imbalance shown earlier, we control for the baseline climate (or the baseline value of the relevant sub-index) for those bureaucrats tracked from baseline. We again find null impacts of the division-level training throughout.

4.2 Ideas Generation and Sharing

How do changes in workplace climate translate into changes in the generation of new ideas by bureaucrats and their willingness to share them? We examine the issue using data from the endline civil servant survey. The results are in Table 5 where we first consider impacts of the classroom training T1 vs. C1.

The first dimension we consider is an index of freedom of expression. The dependent variable is an aggregated 1-4 index of civil servants agreeing/strongly agreeing with the following statements: 'In my work unit, I can freely express my thoughts,' 'In my work unit, expressing your true feelings is welcomed,' 'I proactively report coordination problems in the workplace to the management,' and 'I advise other colleagues against undesirable behaviors that would hamper job performance.' For those randomly assigned to T1, the index of freedom of expression shifts forward by .311 ($p = .030$), corresponding to a 21% increase over the C1 mean. Table A5 shows impacts on each component, showing this result is driven by the ability to express ones feelings ($p = .001$).

We next examine the ability to raise suggestions around new ideas, measured with an aggregated 1-5 index of bureaucrats agreeing/strongly agreeing with the following statements: 'I raise suggestions to improve the unit's working procedure,' 'I make constructive suggestions to improve the unit's operation,' 'I proactively develop and make suggestions for issues that may influence the unit,' 'I proactively suggest new projects which are beneficial to the work unit,' and 'I proactively voice out constructive suggestions that help the unit reach its goals.' Column 2 shows that for T1 this index does not shift overall. Examining individual components in Table A5, we see most do not shift except the ability to make suggestions on procedures ($p = .046$) and on general issues ($p = .059$) – despite the baseline levels of these being very high in C1.

We next consider bureaucrats discussing processes to improve workplace productivity with colleagues. The dependent variable is the proportion of bureaucrats reporting to have monthly or more regular meetings in response to the question, 'Aside from training courses, how often would you say your division discusses ideas or ways to make work processes more efficient?'. In Column 3 of Table 5 we see this shifts forward by 14pp ($p = .029$), corresponding to a 30% increase over the C1 mean. Finally, on the origin of ideas we asked respondents whether individual officers commonly originate ideas. This outcome is shown in Column 4 and shifts forward by 12pp ($p = .055$), corresponding to a 57% increase over the C1 mean.[15]

[15]The samples related to the discussing productivity improvements and origin of ideas are larger because this

The remaining Columns in Table 5 repeat the analysis of ideas for T2 vs. C2. We see that no dimension of ideas generation or sharing improves – in line with there being no shift in workplace climate in T2 divisions. In Table A5 examining sub-components, we see that no component of the freedom of expression index rises. The impacts of T2 on the sub-components of the raising suggestions index are all negative, and we observe significant reductions in the ability to raise suggestions on procedures ($p = .036$). In short, there are some indications that the division-level training actually led to some inhibition of ideas generation in the long term relative to C2 divisions.

4.3 Bureaucratic Performance

The final stage of our analysis examines whether and how shifts in workplace climate, ideas generation and sharing, translate into bureaucratic performance. This helps validate whether the ideas raised and shared are indeed 'good' ideas. We do so using two non-survey based measures: (i) the quality of administrative processes; (ii) task completion by teams.

Administrative Processes We consider administrative processes along three margins: the quality of procedures, adherence to procedures, and the quality of content. The unit of observation is the file, and the sample covers open files (i.e. in use) during or after the training period. Given our research design, we consider files from three types of division: (i) where at least one member was assigned to classroom training T1; (ii) the division was assigned to T2; (iii) the division had no one up for promotion training.

Observations are inverse-weighted by the number of files per division, so that each division receives equal weight. We control for the following file characteristics: whether the assessor stated they had access to all the information needed to assess the file; a dummy stating that there were no challenges encountered in judging the quality of the file; the log number of files in the division; dummies for the file assessors; and the day on which the assessment was done. All specifications control for the number of bureaucrats from the division that attended the ten-day training (in either C1 or T1). Standard errors are clustered at the division-organization level. Divisions that had no one up for promotion training and so have no members in C1 or T1 (or T2) are the omitted category and at the foot of each Column we report their mean outcome.

The results are in Table 6. Column 1 shows that files from divisions with at least one member assigned to T1 have significantly higher quality procedures, the measure shifts forward by .286 ($p = .050$). Column 2 shows files also display higher adherence to procedures, the magnitude of the effect being .322 ($p = .052$). Interestingly, Column 3 indicates that improvement in the quality of procedure does not translate into quality of content of files.

For files from T2 divisions, we see no change in administrative processes, in line with there being no shift in workplace climate or ideas generation or sharing in those divisions.

module was asked in all surveys, while those on the other indices were asked in a subset of surveys.

When comparing the two treatments, the results are not sufficiently precise to observe statistically significant differences between the two.

Task Completion by Divisions Outcomes in the remaining Columns of Table 6 relate to task completion by divisions. Our working sample relates to the completion of 627 tasks at division-level across our core 38 organizations for the endline year 2018 (6-18 months post treatment). We control for the following task characteristics: a dummy stating whether the task is a one-off task or a periodic/regular task; a dummy stating whether the reported task is a single task, or a bundle of tasks; a dummy stating whether the task requires coordination with stakeholders outside the organization; and the log number of tasks in the division. All specifications control for the number of bureaucrats from the division that attended the ten-day training (in either C1 or T1). Observations are inverse-weighted by the number of tasks per division.

The results in Column 4 to 6 examine margins of task completion by divisions. Column 4 shows that there is no change in task initiation under either treatment – although this is not altogether surprising given that the vast majority of assigned tasks do get underway. Column 5 shows that task completion rates (on a 1-5 scale) nudge forward under T1 and nudge down under T2 so that the difference between treatments is borderline significant ($p = .110$). Column 6 shows this effect comes from task completion rates: these rise by .110 for divisions that have at least one individual assigned to T1 ($p = .071$), corresponding to a 68% increase in tasks being fully completed. In contrast, task completion falls slightly in T2 and the difference between treatments is borderline significant ($p = .120$).

4.4 Other Outcomes

To bridge our findings to established literatures on management as a driver of workplace productivity, as well as the more nascent experimental literature on workplace climate, we examine treatment effects on: (i) management practices; (ii) trust and corruption among colleagues; (iii) job satisfaction.

Management Practices Alongside workplace cultures, management practices are another important form of intangible capital within organizations. The importance of top-down management practices for public service delivery in bureaucracies has been documented in high- and low-income contexts [Rasul and Rogger 2018, Bandiera *et al.* 2021, Rasul *et al.* 2021, Englmaier *et al.* 2022]. In our bureaucrat surveys we included modules to measure perceptions of management practices used in a respondents' division. We adapted the Bloom-Sadun-Van Reenen (BVR) management surveys to our setting [Bloom and Van Reenen 2007, Bloom *et al.* 2013, Rasul *et al.* 2021] to measure six dimensions of management: roles, flexibility, incentives, monitoring, staffing and targets, and use these to derive three measures of management practice: (i) the provision of au-

tonomy/flexibility to bureaucrats; (ii) performance related provision of incentives/monitoring to bureaucrats; (iii) a composite index of other practices. We construct a z-score for management practices by division. Table C5 shows the scorecard used to construct each index.[16]

While changing management perceptions was not an explicit part of the new training content, it is plausible such shifts could occur given their focus on bottom-up processes to improve productivity. To be clear, counter to much of the earlier literature, our measures of management practices reflect bottom-up perspectives of bureaucrats on how they are managed (rather than senior's perspectives on the practices they think they are using).

The results are in Table A6. For bureaucrats randomly assigned to T1 relative to those in C1, they view management practices to have improved by $.112\sigma$ ($p = .097$) (Column 1). This is driven by improved perceptions of practices related to performance by $.188\sigma$ ($p = .022$) (Column 2) – partially overlapping with some of the emphasis in the new training module. There is no change in perceptions of practices related to bureaucrat autonomy or other dimensions of management practice. In contrast, we see no changes in perceived management practices for those randomly assigned to division-level training (T2) relative to those in C2.

Trust and Corruption Among Colleagues Changes in workplace climate, including in teamwork climate and junior-senior relations, might naturally alter perceptions of division colleagues. To assess this we asked bureaucrats about their level of trust in colleagues, projects where they were placed under pressure to divert funds, and projects on which they had observed others breaking rules for personal gain. These outcomes are shown in Columns 5 to 7 of Table A6. Neither treatment shifts these outcomes. These findings complement the results of Harris *et al.* [2024], suggesting no free lunch from our innovation-orientated treatment, in that such co-worker ties need to be directly targeted in order to be shifted.[17]

Job Satisfaction If individuals value a workplace climate fostering ideas generation, this can lead them remain in service. We examine this in the remaining Columns of Table A6, asking individuals whether they want to change their job, or whether they agree/strongly agree with the statement, 'Working in the public sector is generally better than working in the private sector.' For both treatments we see null impacts on both measures of job satisfaction. This suggests fostering a

[16] To improve the accuracy of responses and minimize social desirability bias, enumerators asked questions in a style akin to conducting qualitative interviews. They started by asking an open-ended question about how that practice works in the division and asking follow-up questions and requests for illustrative examples. Enumerators used their own judgment to score each practice on a continuous one to five scale, where one corresponds to very poor practice and five corresponds to excellent practice, benchmarking against a pre-defined scoring grid. This approach to measuring management quality has been extensively validated in various contexts [Bloom *et al.* 2012], and our adaptation of it to our context was pre-validated with Ghana's Civil Service [Rasul *et al.* 2021].

[17] Harris *et al.* [2024] experimentally evaluate a training program for police officers in Ghana, designed to foster shared identity and bolster intrinsic motivation to serve. They find the intervention positively affected police officers' values, beliefs regarding on-the-job unethical behavior, reduced officers' propensity to behave unethically as measured by their willingness to cheat in an incentivized game, and improved their attitudes toward citizens.

better workplace climate for ideas has weak implications for staff retention in bureaucracies, even though it leads to more bottom-up ideas being raised and shared. These findings complement Alan *et al.* [2023], again suggesting no free lunch on this dimension from our intervention that targeted a different aspect of workplace climate.[18]

5 Mechanisms

The remainder of our analysis focuses on understanding why the trainings had such contrasting impacts on workplace climate, ideas generation and bureaucratic performance. We do so by first analyzing data from action plans. Specifically, we study the learnings from the training as embodied in the content and implementation of action plans.

The formulation of action plans is central to the new training. Action plans were re-designed as part of our interventions to help bureaucrats conceptualize change and actually apply their new problem solving skills. The plans were produced by trainees on a topic of their choosing, but with guidance, coaching, and feedback from the trainers. Each individual in the new classroom training (T1) developed their own action plan. In the division-level training, the division developed a single action plan, with many teams opting to delegate follow-up of the action plan to a sub-team [Williams and Yecalo-Tecle 2020].[19]

We focus on action plans drafted under T1 and T2 because they have the same format and relate specifically to the new training. We coded action plans from 133 T1 trainees (out of 139 individuals assigned to T1), and from 89 T2 divisions (out of 90 divisions assigned to T2). We present evidence on the quality of action plans, and the ideas proposed in them. To code both dimensions, we worked with civil servants from OHCS, CSTC, and the Management Services Department who had expertise in training, productivity, and management analysis. The coding scheme was developed by the research team, piloted, and adapted in collaboration with these civil servant coders. The coders operated under the supervision of research assistants, and participated in an initial training and follow-up training to ensure consistency of coding, and that there was no influence of treatment assignments on coding practices.

[18] Alan *et al.* [2023] present results from an RCT that aimed to improve workplace climate in terms of leaders behavior and leader-subordinate relationships. They indeed find that such a directly targeted intervention leads to lower separation rates and higher workplace satisfaction.

[19] Such delegation can help ameliorate free-riding in divisions, whereas responsibility was clear for the individual-based training. Of course, this has to be balanced against the fact that divisions agreed their action plans beforehand, and so had more immediate buy-in from all members than plans formulated by individuals under T1. Anecdotally, the impression of trainers and research team members was that delegation of the action plan (as opposed to their collective development) was generally a sign of disinterest by senior team members in using the action plan as an opportunity to make meaningful change.

5.1 Quality of Action Plans

Panel A of Table 7 presents evidence on the quality of action plans developed under T1 and T2. We code quality along five dimensions, each assessed on a 1-5 scale (with a score of 1 indicating the lowest quality). We report the share of action plans that received a coding of 4 or 5.[20]

We start by assessing the single most important dimension of quality: whether action plans actually reflect lessons from the training received. We see that while 35% of action plans under the T1 individual-level training score highly, this falls to 15% for action plans formulated under T2, the difference being significant ($p = .060$). T2 action plans are not coded to be of lower quality than T1 action plans on other dimensions, and are coded to be of higher quality on some.[21]

5.2 Innovations Proposed in Action Plans

Panel B to D of Table 7 examine the ideas suggested in action plans developed under T1 and T2. In each panel, we order aspects of action plans from most to least frequent under T1. Panel B details the kind of work process that action plans focused on. Under each treatment arm, the modal focus is on routine work processes (that at least a third of action plans under each treatment arm cover), but T2 action plans are slightly less likely to focus on such routine processes ($p = .089$). Across T1 and T2, action plans are equally likely to focus on other core work tasks related to'data, monitoring and oversight', 'communications', 'personnel', and 'financial management'. However, a key distinction is that ideas proposed in T2 are significantly more likely to focus on 'other

[20]The lessons from training outcome is derived from coders' responses to, *Has the individual/ division used lessons from the training in coming up with the Action Plan?* Coders were asked to rate this on a scale of 1-5, with scores benchmarked as follows: 1 = No explicit or implicit connection between Action Plan and training content; 3 = Some lessons are apparent or explicitly identified, but they are only somewhat relevant to the Action Plan; 5 = The Action Plan draws heavily on training ideas and content. The level of detail outcome is derived from coders' responses to, *In your view, how thorough and detailed is this Action Plan (compared to others of the same format)?* Coders were asked to rate this on a scale of 1-5, with scores benchmarked as follows: 1 = Hastily written and not detailed; 3 = Some detail, but not all sections are completed thoroughly; 5 = All sections are thoroughly completed. The overall feasibility outcome is derived from coders' responses to, *In your view, how likely is this Action Plan to actually be implementable?*" Coders were asked to rate this on a scale of 1-5, with scores benchmarked as follows: 1 = Not at all feasible – has unrealistic assumptions or requirements; 3 = Somewhat feasible – there will be significant challenges, but the plan has a realistic way to overcome them; 5 = Very feasible – plan convincingly shows how it can overcome all obstacles. The degree of ambition outcome is derived from coders' responses to, *In your view, how ambitious is this Action Plan?* Coders were asked to rate this on a scale of 1-5, with scores benchmarked as follows: 1 = Not ambitious – reform is minor and will not make much difference on the individual's or division's productivity; 3 = Somewhat ambitious – improvement would be significant but not transformative; 5 = Improvement would be transformative to the individual or division if implemented successfully. Finally, the success fully defined outcome is derived from coders' responses to, *Is there a clear definition of what success would be for this improvement?* Coders were asked to code this categorically, with categories defined as follows: 1 = Yes, success is explicitly or implicitly defined clearly, 2 = Somewhat, 3 = Not at all. Here we report the share of action plans that received a coding of 3.

[21]To see whether the lack of training content being reflected in action plans relates to some basic misunderstanding, we can also examine how action plans described who was responsible for implementing the idea. In line with the design of the treatments, T2 action plans are significantly more likely to involve the involvement of colleagues from the division (73% versus 51% of those in T1, $p = .001$), and significantly less likely to involve a single individual (3% versus 12% in T1, $p = .024$).

administrative' tasks than T1 plans ($p = .004$).

Panel C shows how the nature of the proposed innovation differs from current practices. Under each treatment arm, the modal focus is on implementing existing practices that were not actually being executed (this might reflect why only around a quarter of action plans under T1 or T2 were viewed to be ambitious). Around 15% of action plans proposed a truly innovative practice, and this does not differ between treatment arms.

Panel D details the implementation activities envisioned. The key distinctions between treatments is that T2 action plans are: (i) significantly less likely to plan further meetings within the division (34% vs 46%, $p = .072$); (ii) significantly more likely to need the division to acquire more resources (either financial or time) (38% vs 27%, $p = .081$); (iii) significantly more likely to envisage the provision of further training (30% vs 20%, $p = .065$); (iv) significantly more likely to need division members to meet with actors outside the organization (16% vs 7%, $p = .032$). These differences start to highlight why T2 might have been ineffective: if hierarchical norms lead to a strong pre-existing consensus around what kinds of action are needed to overcome productivity constraints and bottlenecks, that can lead divisions to fall back into group think towards the kinds of system-wide innovations highlighted in the earlier qualitative evidence (Table B2). These aim for unrealistic system-wide changes in operations, less related to routine or other core tasks, and typically involving managing engagement with multiple other agencies and stakeholders outside the ministry.[22]

5.3 Implementation of Action Plans

We consider the implementation of action plans from two perspectives. Panel A of Table 8 reports bureaucrats *ex ante* perceived obstacles to implementation, as recorded when action plans were originally drafted. Here we observe no statistically difference in perceived future obstacles between T1 and T2. Panel B then moves forward in time to the endline bureaucrat survey, where individuals report *ex post* obstacles to implementation. Differences in perceived implementation barriers between T1 and T2 emerge *ex post*. Specifically, bureaucrats in T2 are significantly more likely to report others in their division not being supportive ($p = .034$), and a lack of resources or other logistical constraints ($p = .034$) – in line with the nature of action plan proposals described before being focused around system-wide changes in operations. The fact that division colleagues were reported to be less supportive under T2 also goes against one of the original motivations for

[22]The focus of action plan outcome is derived from coders' responses to, *What is the main aspect of their work have they identified for productivity improvement in the Action Plan?* Coders were asked to select one option that was most applicable. The nature of proposed change outcome is derived from coders' responses to, *What type of change have they proposed in this area?*" Coders selected one option that was most applicable. The main activities envisioned outcome is derived from coders' responses to, *What types of steps were identified to implement this improvement?* Coders could select all applicable options. The implementation responsibility outcome is derived from coders' responses to, *Which other colleagues/stakeholders do they identify as important for collaboration?* Coders could select all applicable options. The obstacles envisioned outcome is derived from coders' responses to, *What difficulties have they foreseen in implementing this improvement?* Coders could select all applicable options.

the division level training – further supporting the idea that the division-based training format backfired in trying to change workplace climate.[23]

Does all this mean that T2 is ineffective because the intended innovation was too unrealistic – reverting back to pre-existing norms of focusing on system-wide rather than incremental innovations, or because steps were never actually taken to implement T2 action plans? Panel C distinguishes these explanations by examining endline survey data on reports on steps taken to implement action plans: here we see little difference between T1 and T2, except that under T2, individuals work alone in having to implement their plans, despite their top-down systems wide proposed change.[24]

5.4 Presence of Directors

A key theme of the qualitative evidence on workplace culture was the nature of hierarchical relationships, that often caused ideas not to be raised or follow through on (Table B1). Hence it is natural to consider whether the ineffectiveness of T2 can be explained by the presence of Directors at the division-level training because it leads to: (i) an inhibition of open discussion of issues and potential improvements to work processes in the division; (ii) among ideas that are raised, an inherent conservatism in their innovativeness.[25]

Qualitative Evidence In our qualitative interviews we prompted respondents to discuss how they felt about their Director being present at the division-level training. Interview extracts are presented in Table B4. These reflect both negative and positive views of the presence of Directors. Negative views reflect the inhibition of discussion, in part because of concerns over how superiors would receive new ideas. On the other hand, the positive views expressed included that Directors need to be present at the initiation of any discussion around innovation, or because Directors can themselves help initiate new ideas and discussion.

[23] In Panels B and C based on endline data rather than the action plans, when testing for differences between T1 and T2 we control for gender, education (having an undergraduate degree or not), whether the individual works for a human resource management division or other division, and randomization strata.

[24] We can validate whether this distinction between simple incremental ideas being pursued under T1 and more system-wide changes being pursued under T2 is reflected in how the organization of work within divisions changes at endline. We do so in terms of how tasks are assigned, whether work is conducted with colleagues, and the total time spent at work. The results are in Table A7. Panel A shows that for T1 relative to C1, these aspects of the work environment do not shift. Hence improvements in workplace climate and ideas generation occur even when there is no change in how tasks are assigned by seniors, or total time spent at work. Panel B shows that bureaucrats in T2 divisions significantly reduce the percentage of tasks they conduct with division colleagues: the magnitude of the effect is 7.6pp, relative to those in C2 that report two thirds of tasks are conducted with colleagues ($p = .029$). This is in line with the evidence from action plans and their implementation, whereby the innovations proposed in action plans relate more to working with outside stakeholders.

[25] Managers matter for public sector productivity [Fenezia 2022, Kala 2024]. Such 'manager fixed effects' results can be microfounded through models of managerial attention [Dessein and Santos 2021]. This is distinct from the reason we put forward – that managers shape how open workplace climates are to bottom-up innovation.

Quantitative Evidence Given the mixed qualitative evidence, we examine the issue quantitatively by considering how our core results for T2 vary with the presence of division Directors during training. In the 93 divisions assigned to T2, Directors were present in 57% of them. We estimate specifications analogous to (1) where we allow for the effects of T2 to differ depending on the presence of the division Director (we cannot estimate the direct effect of the presence of Directors as this is not defined among C2 divisions).

Overall, we continue to find largely null impacts of T2 irrespective of the presence of Directors. Specifically, we find no differential impacts on workplace climate up to 18 months after training takes place – either on the aggregate index or nearly each sub-index except for perceptions of management (Table A9). We also find no evidence that the presence or not of Directors at T2 training has lasting impacts on ideas generation, sharing or origination (Table A10), or most margins of public service delivery (Table A11).

Finally, Table A12 shows how the quality and content of action plans drawn up in T2 varies with the presence of Directors. We do not see substantive differences along margins related to the quality, focus or nature of innovation proposed, *ex ante* or *ex post* perceived obstacles, or steps taken to implement the action plan (Panels A to C, E to G). However, in the presence of Directors action plans are significantly more likely to detail ideas that require additional resource or time inputs ($p = .046$) (Panel D). Taken together, this evidence suggests the hierarchical culture of bureaucracies is important because it entrenches attitudes to bottom-up innovation, not because the mere presence of Directors at division-level training stymies discussion, ideas proposed or steps taken to implement them.

5.5 Other Mechanisms

We consider two additional mechanisms that *a priori* could explain the differential effectiveness of individual and division-level training: (i) spillovers from individual trainees from T1 driving changes in work processes among colleagues; (ii) heterogeneity within divisions making it harder to change workplace climate and increasing openness to innovation and new ideas. We relegate discussion of each alternative mechanism to the Appendix because the bottom line in each case we do not find robust quantitative evidence that they help explain the core results on workplace climate, ideas generation and sharing, and bureaucratic performance.

5.6 Summary

Bringing the results together paints a coherent picture: division-level training is less effective than individual training in shifting workplace climates, ideas generation and sharing because the action plans drawn up by divisions collectively fail to integrate in the key ideas of the novel training provided and focus on incremental ideas rather than system-wide change. As a result, they focus

on non-core work processes, suggest fewer follow-up meetings to implement ideas, and are less realistic and harder to achieve in that they require acquiring additional resources, the provision of further training, and have a greater need to meet with actors outside the organization. In short, division-level action plans in T2 much more reflect pre-existing workplace cultures that we documented using our qualitative data, where training was often seen to be ineffective because constraints on productivity were due to service-wide inefficiencies, and thus a lack of resources was cited being the single largest obstacle to addressing performance issues. This is very far from the spirit of the novel training provided (and internalized under T1) where individuals/divisions were encouraged to break down complex problems into manageable solutions.

Our results echo findings from other settings suggesting the impact of interventions in organizations can depend critically on pre-existing workplace cultures – where we highlight novel insights arising from the fact that organizations are hierarchical [Blader *et al.* 2020, Celhay *et al.* 2024]. These are widely applicable lessons for other contexts.

Two further points are of note.

First, our results go further than just showing that pre-existing workplace cultures overwhelm the emergence of new ideas in action plans arising from division-level training. It could have been the case that the division-level training changed workplace climates, ideas generation and sharing *outside* of the specific action plans developed. This does not occur – the very nature of the group level intervention seems to dull learnings from the training being applied more widely day-to-day.

Second, given our design, it is natural to ask whether individual and division-level treatments interact. We examine this in Table A8 for workplace climate, revealing two main insights. First, it remains the case that T1 significantly shifts forward the aggregate index of workplace climate, and sub-indices related to performance, and stress recognition. T2 remains largely ineffective, and the impacts of T1 are significantly greater than for T2 on the aggregate index of culture ($p = .025$), as well as sub-indices related to performance climate ($p = .015$), fostering new ideas ($p = .085$), and stress recognition ($p = .083$). Second, there is little robust evidence of interaction effects between the treatments on either the aggregate index of division culture, or sub-indices with the exception of stress recognition. However most of the interaction point estimates are negative, so the positive impacts of individual training are weakened if the individual returns to a division that is treated.

6 Final Remarks

State capacity is important for economic development, helping spur growth and poverty reduction [Acemoglu 2005, Besley and Persson 2010]. Yet building state capacity has proven to be hard in much of Sub-Saharan Africa [Andrews *et al.* 2017]. This is not for a lack of intervention: 25% of overseas development assistance in Africa has been targeted towards capacity building in bureaucracies [Adamolekun *et al.* 2016]. Much of this has taken the form of training to improve the

human capital of bureaucrats. We complement such approaches using the lens of organizational economics to focus on the value of changing workplace cultures that govern relations between bureaucrats. We do so because bureaucratic organizations can be hierarchical, an important consequence of which is to stymie innovation, and something that can be offset through individual but not division-level intervention.

Given the inherent difficulty of valuing bureaucratic output, we make no attempt to measure the cost effectiveness of the new individual training. However we reiterate that task completion rates rise by 11pp (68%) for divisions that have at least one individual assigned to T1 ($p = .071$), suggesting non trivial benefits given the aggregate budget of ministries we worked with comprise more than $5bn. Moreover, our study findings were incorporated by the OHCS into their training plans: (i) the T1 curriculum was integrated into the permanent promotion training curriculum for senior civil servants; (ii) they paused plans to roll out division-based training more widely.

We make three final remarks on a future research agenda stemming from our work.

First, our findings hinge on two key features of our context: contractual inefficiencies abound [Dixit 2002, Besley and Ghatak 2005], and organizations are hierarchical, both of which leave scope for workplace culture to be a key driver of individual and collective behavior [O'Reilly 1989, Kreps 1990, Baron and Kreps 2013]. All organizations – in public and private sectors – are subject to both features to some extent, providing a four-way classification for future work to explore.

Second, our approach to understanding soft interventions that aim to shift workplace cultures combines qualitative and quantitative data. Given the plethora of definitions of workplace culture, we see value in mixed-methods approaches for future work. This is especially so given that the importance of relational contracts within organizations has long been recognized, but remains hard to measure. As Gibbons and Henderson [2012] describe, these contracts can mesh together to form workplace culture. Such informal relations between senior and juniors can help explain variations in performance in seemingly otherwise similar organizations, depending on their ability to solve problems of credibility and clarity that underlie effective relational contracts.

Finally, a bottom line from our work is that workplace cultures can be hard to shift, even when interventions seek to shift them through group-level coordination. This has important implications for thinking of organizational culture as equilibrium selection devices [Chassang 2010, Baron and Kreps 2013] and related work on group think [Benabou 2013]. Our work leaves open questions about the role that leaders play in determining workplace cultures, but this seems to be important given evidence highlighting the value of leaders/managers through top-down practices [Bloom and Van Reenen 2007, Fenezia 2022, Kala 2024], shaping workplace climate to reduce harassment [Alan *et al.* 2024], exploiting social connections to drive productivity [Bandiera *et al.* 2009], or coordinating workers [Boudreau *et al.* 2024]. We also see high value in future research that integrates such ideas with the literature on the psychological basis for managerial aversion to employee voice to better understand how hierarchical cultures might emerge in the first place [Edmondson 1999, Ashford *et al.* 2009, Fast *et al.* 2014, Morrison 2014].

A Appendix

A.1 Other Mechanisms

A.1.1 Spillovers

The qualitative evidence on workplace cultures emphasized the importance of relations between colleagues in being able to raise new ideas and have them adopted (Table B1). A body of work in economics has examined the impact of managers on subordinates [Bandiera *et al.* 2007, 2009, Hoffman and Tadelis 2021] but much less is known about spillovers from lower-tier to higher-tier workers – an exception is Espinosa and Stanton [2022] who study spillovers from training workers on managers productivity in government agencies in Columbia, where the key mechanism is a reduced need for managers need to help lower level employees.

We pursue this idea of spillovers by examining how the presence of T1 trainees impact co-workers. We start by examining spillovers to those that were not subject to any form of training because they were not due to attend training in our study period as part of their promotion cycle. We refer to this group of bureaucrats as C0. In Table A13 we show that on the aggregate index of workplace climate and nearly all sub-components, C0 bureaucrats are not differentially impacted in terms of their perceptions of workplace climate depending on whether they have a colleague in their division that received the individual T1 training. Table A14 confirms similarly null impacts when examining spillovers of T1 in terms of ideas generation, sharing or origination among C0 bureaucrats.

A.1.2 Heterogeneity Within Divisions

Finally, we consider the possibility that because divisions can comprise a diverse set of individuals, this makes it harder to drive forward changes in workplace climate and openness to innovation and new ideas. As Charness and Greico [2023] discuss, the literature in organizational studies has consistently highlighted group-based innovation to be facilitated by members' trust and cooperation [Tushman and O'Reilly 1997, Anderson and West 1998], shared or interdependent objectives [West 1990, Hülsheger *et al.* 2009, Guiso *et al.* 2015].

We pursue such intuitions by examining diversity arising from two sources: workplace relations, and broader views of public service. On workplace relations, we use our baseline bureaucrat survey to construct measures of diversity in divisions based on: (i) the extent to which individuals are day-to-day engaged in different types of task to their division co-workers; (ii) how much bureaucrats reports trusting division colleagues. On broader views of public service, we measure the diversity of divisions using baseline measures of: (i) how much bureaucrats perceive the service accepts low-levels of corruption as a pragmatic response to delivering public services; (ii) Public Service Motivation scores [Perry 1996]. We estimate heterogeneous effects of the division-level training

T2 in our core results depending on these forms of division heterogeneity.[26]

The bottom line is that we find little evidence that heterogeneity within divisions impacts treatment effects of T2 on changes in workplace climate, ideas generation/sharing, or public service delivery. More precisely, we generally find few differential impacts of any of these dimensions on the aggregate index of workplace climate (Tables A15A, A15B). We find that in divisions where at baseline bureaucrats have a greater acceptance of low-level corruption, bureaucrats are significantly less likely to report having the freedom to express ideas or to raise suggestions (Table A16). Finally, we find limited impacts of any of dimension of heterogeneity interacting with T2 in terms of bureaucratic effectiveness (Tables A17A, A17B), although in divisions with greater levels of trust among co-workers, T2 leads to significantly higher task completion rates.[27]

References

[1] ACEMOGLU.D (2005) "Politics and Economics in Weak and Strong States", *Journal of Monetary Economics* 52: 1199-226.

[2] ADAMOLEKUN ET AL. (2016) *Building State Capacity in Africa: New Approaches, Emerging Lessons*, World Bank Group.

[3] ADHVARYU.A, J-F.GAUTHIER, A.NYSHADHAM AND J.TAMAYO (2024) "Absenteeism, Productivity, and Relational Contracts Inside the Firm," *Journal of the European Economic Association* 22: 1628-77.

[4] AGHION.P AND P.HOWITT (1992) "A Model of Growth Through Creative Destruction," *Econometrica* 60: 323-51.

[5] AGHION.P AND P.HOWITT (1998) *Endogenous Growth Theory*, MIT Press.

[26] The task-based measure of heterogeneity is constructed from individual responses to the question, *Which of the following processes are you most closely involved with?* They could select multiple types of processes (such as writing reports, answering correspondence, and responding to audit queries) from a showcard. From this we constructed a Herfindahl index, aggregated to the division level, on the share of bureaucrats working on different processes in the division at baseline. The trust-based measure of heterogeneity is based on question where we asked respondents, *On a scale of 1 to 4, how much do you trust each of the officers in your division?* We then created an index using the share of bureaucrats that reported as score of 4. The acceptance of corruption measure is the share of bureaucrats in the division that agreed with the statement, *The government regards low-level corruption as unfortunate but unavoidable until it can pay better salaries.* Finally, the Public Service Motivation measure uses the Perry [1996] sub-indices related to attraction to policy making, compassion for underprivileged people, commitment to the public interest, and self-sacrifice.

[27] We also examined heterogeneity by division size: we find little evidence that impacts on workplace climate, administrative or bureaucratic performance vary under T2 with division size. However, the one margin we find precise effects are in relation to the freedom of expression index. We find that in larger divisions individuals report being less free to express ideas or raise suggestions, but these negative effects are mitigated under T2 in larger divisions. This goes against the notion that the effects of T2 are dulled because of collective action problems in engaging with training about ideas generation.

[6] AGHION.P AND J.TIROLE (1997) "Formal and Real Authority in Organizations," *Journal of Political Economy* 105: 1-29.

[7] AHMED.P (1998) "Culture and Climate for Innovation," *European Journal of Innovation Management* 1: 30-43.

[8] AKCIGIT.U AND J.VAN REENEN (2023) *The Economics of Creative Destruction*, Harvard Univerity Press.

[9] ALAN.S, G.COREKCIOGLU AND M.SUTTER (2023) "Improving Workplace Climate in Large Corporations: A Clustered Randomized Intervention," *Quarterly Journal of Economics* 138: 151-203.

[10] AMABILE.T.M (1996) *Creativity in Context: Update to the Social Psychology of Creativity*, Westview Press.

[11] ANDERSON.M.L (2008) "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association* 103: 1481-95.

[12] ANDERSON.N.R AND M.A.WEST (1998) "Measuring Climate for Work Group Innovation: Development and Validation of the Team Climate Inventory," *Journal of Organizational Behavior* 19: 235-58.

[13] ANDREWS.M, L.PRITCHETT AND M.WOOLCOCK (2017) *Building State Capability: Evidence, Analysis, Action*, OUP.

[14] ASHFORD.S, K.SUTCLIFFE AND M.CHRISTIANSON (2009) "Speaking Up and Speaking Out: The Leadership Dynamics of Voice in Organizations," in J.Greenberg and M.Edwards (eds.) *Voice and Silence in Organizations*, Emerald Group Publishing.

[15] ASHRAF.N AND O.BANDIERA (2018) "Social Incentives in Organizations," *Annual Review of Economics* 10: 439-63.

[16] BANDIERA.O, I.BARANKAY AND I.RASUL (2007) "Incentives for Managers and Inequality Among Workers: Evidence From a Firm Level Experiment," *Quarterly Journal of Economics* 122: 729-75.

[17] BANDIERA.O, I.BARANKAY AND I.RASUL (2009) "Social Connections and Incentives: Evidence From Personnel Data," *Econometrica* 77: 1047-94.

[18] BANDIERA.O, M.C.BEST, A.Q.KHAN AND A.PRAT (2021) "The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats," *Quarterly Journal of Economics* 136: 2195-242.

[19] BARON.J.N AND D.M.KREPS (2013) "Employment as an Economic and a Social Relationship,", in R.Gibbons and J.Roberts (eds.) *The Handbook of Organizational Economics*, Princeton University Press.

[20] BENABOU.R (2013) "Groupthink: Collective Delusions in Organizations and Markets," *Review of Economic Studies* 80: 429-62.

[21] BESLEY.T, R.BURGESS, A.Q.KHAN AND G.XU (2021) "Bureaucracy and Development," *Annual Review of Economics* 14: 397-424.

[22] BESLEY.T AND M.GHATAK (2005) "Competition and Incentives with Motivated Agents," *American Economic Review* 95: 616-36.

[23] BESLEY.T AND T.PERSSON (2010) "State Capacity, Conflict and Development," *Econometrica* 78: 1-34.

[24] BESLEY.T AND T.PERSSON (2024) "Organizational Dynamics: Culture, Design, and Performance," *Journal of Law, Economics, and Organization* 40: 394-415.

[25] BLADER.S.L, C.GARTENBERG, R.HENDERSON AND A.PRAT (2015) "The Real Effects of Relational Contracts," *AER: Papers and Proceedings* 105: 452-6.

[26] BLADER.S.L, A.PRAT AND C.GARTENBERG (2020) "The Contingent Effect of Management Practices," *Review of Economic Studies* 87: 721-49.

[27] BLOOM.N, B.EIFERT, A.MAHAJAN, D.MCKENZIE AND J.ROBERTS (2013) "Does Management Matter? Evidence from India," *Quarterly Journal of Economics* 128: 1-51.

[28] BLOOM.N, C.I.JONES, J.VAN REENEN AND M.WEBB (2020) "Are Ideas Getting Harder to Find?" *American Economic Review* 110: 1104-44.

[29] BLOOM.N, R.SADUN AND J.VAN REENEN (2012) "The Organization of Firms Across Countries," *Quarterly Journal of Economics* 127: 1663-1705.

[30] BLOOM.N AND J.VAN REENEN (2007) "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics* 122: 1351-1408.

[31] BOLTON.P, M.BRUNNERMEIER AND L.VELDKAMP (2013) "Leadership, Coordination and Corporate Culture," *Review of Economic Studies* 80: 512-37.

[32] BOUDREAU.L, R.MACCHIAVELLO, V.MINNI AND M.TANAKA (2024) "Leaders in Social Movements: Evidence from Unions in Myanmar," *American Economic Review* forthcoming.

[33] CELHAY.P.A, P.GERTLER, M.OLIVARES AND R.UNDURRAGA (2024) How Managers Can Use Purchaser Performance Information to Improve Procurement Efficiency, NBER WP32141.

[34] CHARNESS.G AND D.GRIECO (2023) "Creativity and Corporate Culture," *Economic Journal* 133: 1846-70.

[35] CHASSANG.S (2010) "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts," *American Economic Review* 100: 448-65.

[36] CREMER.J (1993) "Corporate Culture and Shared Knowledge," *Industrial and Corporate Change* 2: 351-86.

[37] DESSEIN.W AND T.SANTOS (2021) "Managerial Style and Attention," *AEJ: Microeconomics* 13: 372-403.

[38] DIXIT.A (2002) "Incentives and Organizations in the Public Sector: An Interpretive Review," *Journal of Human Resources* 37: 696-727.

[39] EDMONDSON.A (1999) "Psychological Safety and Learning Behavior in Work Teams," *Administrative Science Quarterly* 44: 350-83.

[40] EDMONDSON.A AND A.LEI (2014) "Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct," *Annual Review of Organizational Psychology and Organizational Behavior* 1: 23-43.

[41] ENGLMAIER.F, G.MUHLHEUSSER, A.ROIDER AND N.WALLMEIER (2022) Management and Performance in the Public Sector: Evidence from German Municipalities, mimeo LMU.

[42] ESPINOSA.M AND C.T.STANTON (2022) Training, Communications Patterns, and Spillovers Inside Organizations, NBER WP30224.

[43] FAST.N, E.BURRIS AND C.BARTEL (2014) "Managing to Stay in the Dark: Managerial Self-Efficacy, Ego Defensiveness, and the Aversion to Employee Voice," *Academy of Management Journal* 57: 1013-34.

[44] FENEZIA.A (2022) "Managers and Productivity in the Public Sector," *Econometrica* 90: 1063-84.

[45] FERNANDEZ.S AND T.MOLDOGAZIEV (2012) "Using Employee Empowerment to Encourage Innovative Behavior in the Public Sector," *Journal of Public Administration Research and Theory* 23: 155-87.

[46] FINAN.F, B.A.OLKEN AND R.PANDE (2017) "The Personnel Economics of the State," in A.V.Banerjee and E.Duflo (eds.) *Handbook of Economic Field Experiments* Vol. 2.

[47] GARTENBERG.C, A.PRAT AND G.SERAFEIM (2019) "Corporate Purpose and Financial Performance," *Organization Science* 30: 1-18.

[48] GIBBONS.R AND R.HENDERSON (2012) "Relational Contracts and Organizational Capabilities," *Organization Science* 23: 1350-64.

[49] GROSS.D.P (2020) "Creativity Under Fire: The Effects of Competition on Creative Production," *Review of Economics and Statistics* 102: 583-99.

[50] GUISO.L, P.SAPIENZA AND L.ZINGALES (2015) "Corporate Culture, Societal Culture, and Institutions," *American Economic Review* 105: 336-9.

[51] HARRIS.D, O.BORCAN, D.SERRA, H.TELLI, B.SCHETTINI AND S.DERCON (2024) Proud to Belong: The Impact of Ethics Training on Police Officers in Ghana, IZA DP17006.

[52] HASSAN.S (2015) "The Importance of Ethical Leadership and Personal Control in Promoting Improvement-Centered Voice among Government Employees," *Journal of Public Administration Research and Theory* 25: 697-719.

[53] HEINING.J, S.JAEGER AND B.SCHOEFER (2021) "Labor in the Boardroom," *Quarterly Journal of Economics* 136: 669-725.

[54] HERMALIN.B.E (1998) "Toward an Economic Theory of Leadership," *American Economic Review* 88: 1188-206.

[55] HERMALIN.B.E (2001) "Economics and Corporate Culture," in C.L.Cooper, S.Cartwright and P.C.Earley (eds.) *International Handbook of Organizational Culture and Climate*, Wiley.

[56] HOFFMAN.M AND S.TADELIS (2021) "People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis," *Journal of Political Economy* 129: 243-85.

[57] JASSAWALLA.A.R AND H.C.SASHITTAL (2002) "Cultures that Support Product Innovation Processes," *Academy of Management Executive* 16: 42-53.

[58] JONES.C.I (2019) "Paul Romer: Ideas, Nonrivalry, and Endogenous Growth," *Scandinavian Journal of Economics* 121: 859-83.

[59] KALA.N (2024) "The Impacts of Managerial Autonomy on Firm Outcomes," *Econometrica* 92: 1777-800.

[60] KAUFMANN.DM A.KRAAY AND M.MASTRUZZI (2010) The Worldwide Governance Indicators: Methodology and Analytical Issues, World Bank Policy Research WP5430.

[61] KREPS.D.M (1990) "Corporate Culture and Economic Theory," in J.E.Alt and K.A.Shepsle (eds.) *Perspectives on Positive Political Economy*, CUP.

[62] MARTINEZ.E.A, N.BEAULIEU, R.GIBBONS, P.PRONOVOST AND T.WANG (2015) "Organizational Culture and Performance," *AER: Papers and Proceedings* 105: 331-5.

[63] MORRISON.E (2014) "Employee Voice and Silence," *Annual Review of Organizational Psychology and Organizational Behavior* 1: 173-97.

[64] MURALIDHARAN.K AND A.SINGH (2023) Improving Public Sector Management at Scale: Experimental Evidence on School Governance in India, mimeo UCSD.

[65] O'REILLY.C.A. (1989) "Corporations, Culture, and Commitment: Motivation and Social Control in Organizations," *California Management Review* 31: 9-25.

[66] PERRY.J.L (1996) "Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity," *Journal of Public Administration Research and Theory* 6: 5-22.

[67] PHELPS.E.S (2013) *Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge, and Change*, Princeton University Press.

[68] RASUL.I AND D.ROGGER 2018 "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *Economic Journal* 128: 413-46.

[69] RASUL.I, D.ROGGER AND M.J.WILLIAMS (2021) "Management, Organizational Performance, and Task Clarity: Evidence from Ghana's Civil Service," *Journal of Public Administration Research and Theory* 31: 259-77.

[70] ROMER.P.M (1990) "Endogenous Technological Change," *Journal of Political Economy* 98: S71-102.

[71] ROTEMBERG.J.J AND G.SALONER (2000) "Visionaries, Managers, and Strategic Direction," *RAND Journal of Economics* 31: 693-716.

[72] SCHEIN.E.H (1985) *Organizational Culture and Leadership*, Jossey-Bass.

[73] SEXTON.J.B ET AL. (2006) "The Safety Attitudes Questionnaire: Psychometric Properties, Benchmarking Data, and Emerging Research," *BMC Health Services Research* 6:44.

[74] TUSHMAN.M.L AND C.A.O'REILLY III (1997) Winning Through Innovation: A Practical Guide to Leading Organizational Change and Renewal, Harvard Business School Press.

[75] VAN DEN STEEN.E (2005) "Organizational Beliefs and Managerial Vision," *Journal of Law, Economics, & Organization* 21: 256-83.

[76] WEBER.R.A AND C.F.CAMERER (2003) "Cultural Conflict and Merger Failure: An Experimental Approach," *Management Science* 49: 400-15.

[77] WEST.M.A (1990) "The Social Psychology of Innovation in Groups," in M.A.West and J.L.Farr (eds.) *Innovation and Creativity at Work: Psychological and Organizational Strategies*, Wiley.

[78] WILLIAMS.M.J AND L.YECALO-TECLE (2020) "Innovation, Voice, and Hierarchy in the Public Sector: Evidence from Ghana's Civil Service," *Governance* 33: 789-807.

[79] WILLIAMS.M.J (2025) *Reform as Process: Implementing Change in Public Bureaucracies*, forthcoming.

[80] WILSON.J (1989) *Bureaucracy: What Government Agencies Do and Why They Do It*, Basic Books.

Table 1: Workplace Climate

Sub-index	Question Wording	Mean	SD Overall	SD Within Organizations	SD Between Organizations
Teamwork climate (1-5)	It is easy for personnel here to ask questions when there is something that they do not understand	3.96	.722	.605	.567
	Disagreements in this division are resolved appropriately (i.e., not who is right, but what is best for the service)				
	The managers and other officers here work together as a well-coordinated team				
	Communication breakdowns that lead to delays in delivery of services are common				
Performance climate (1-5)	The culture in this division makes it easy to learn from the errors of others	4.24	.656	.549	.501
	You know the proper channels to direct questions regarding correct bureaucratic process in this division				
	You receive appropriate feedback about your performance				
	Bureaucratic errors are handled appropriately in this division				
	You are encouraged by your colleagues to report any work concerns you may have				
	You would feel happy being served as a Ghanaian citizen by this division				
Stress recognition (1-5)	Fatigue impairs your performance during high-pressure situations (e.g. when there are heavy demands on your division)	2.39	1.37	1.18	.972
Perceptions of management (1-5)	Staff (divisional management) doesn't knowingly compromise division services	3.71	.723	.618	.530
	Staff (divisional management) supports your daily efforts				
	You get adequate, timely info about events that might affect your work from your division.				
	Your performance has no influence on your career progression				
Working conditions (1-5)	All the necessary information for diagnostic and effective decision making is routinely available to you	3.97	.939	.805	.657
	Trainees in your division are adequately supervised				
Fostering new ideas (1-5)	Your suggestions about work place productivity would be acted upon if you expressed them to management	4.27	.635	.533	.486
	Staff (divisional management) in this division are quick to adopt (are open to) new ways of doing things.				
	You can see lots of ways to make your division work better.				
Aggregate Index (1-5)		3.96	.513	.423	.400

Notes: Performance climate is referred to as safety climate in Sexton et al. [2006]. Text in (parentheses) indicates item phrasing administered to non-management-level staff; all other items are identical for all respondents. We report division-level statistics for baseline indicators of workplace climate using raw scores. The mean is for the overall index across all divisions. The standard deviation shown is for the overall index across all divisions, within organizations, and between organizations.

Table 2: Balance, Classroom Training

Means, standard deviations in parentheses, p-values on tests of equality in brackets

	(1) Standard classroom training (C1)	(2) New classroom training (T1)	(3) p-value [C1=T1]	(4) Not up for training (C0)	(5) p-value [C0=C1]
Panel A: Bureaucrat Characteristics					
Gender [female=1]	.500	.712	[.001]	.486	[.145]
Age	38.4	37.7	[.411]	41.8	[.035]
	(1.98)	(2.22)		(.240)	
Undergraduate degree [yes=1]	.875	.916	[.314]	.765	[.942]
Postgraduate degree [yes=1]	.125	.141	[.820]	.339	[.622]
Years in the civil service	10.6	9.80	[.309]	13.91	[.001]
	(2.97)	(3.06)		(.279)	
Panel B: Division Characteristics					
Working in Finance and Administration [yes=1]	.375	.434	[.363]	.453	[.528]
Working in Policy, Planning, Monitoring and Evaluation [yes=1]	.500	.468	[.599]	.327	[.092]
Working in Human Resources [yes=1]	.125	.077	[.349]	.053	[.010]
Working in Research and Statistics [yes=1]	.000	.033	[.552]	.110	[.967]
Working in other divisions [yes=1]	.036	.030	[.449]	.058	[.073]
Panel C: Workplace Climate and Ideas					
Aggregate Climate Index	-.104	-.018	[.511]	-.036	[.649]
	(.149)	(.071)		(.019)	
Fostering New Ideas	-.364	-.437	[.678]	-.040	[.622]
	(.588)	(.614)		(.027)	

Notes: These statistics are based on those matched from the training interventions through to the endline survey of bureaucrats. In Panel A, we report individual covariates as measured in the endline civil servants survey. In Panel B, we report division-level characteristics. In Panel C we report the aggregate workplace climate z-score and the sub-index z-score for fostering new ideas based on baseline data. In Columns 3 and 5, the p-values are calculated from an underlying regression that controls for the randomization strata for the C1/T1 intervention, and standard errors are clustered at the division-organization level.

Table 3: Balance, Division Training

Means, standard deviations in parentheses
p-values on tests of equality in brackets

	(1) Control division (C2)	(2) Treated division (T2)	(3) p-value [C2=T2]
Panel A: Division Characteristics			
Number of bureaucrats in division	6.94	9.36	[.079]
	(1.07)	(.858)	
Share of women in division	.459	.555	[.036]
Average age	42.1	41.7	[.698]
	(.785)	(.475)	
Share of bureaucrats with undergraduate degree	.858	.800	[.095]
Share of bureaucrats with postgraduate degree	.404	.342	[.198]
Average years in the civil service	12.8	13.8	[.298]
	(.783)	(.551)	
Finance and Administration [yes=1]	.419	.333	[.284]
Policy, Planning, Monitoring and Evaluation [yes=1]	.306	.289	[.817]
Research and Statistics [yes=1]	.129	.200	[.238]
Human Resources [yes=1]	.081	.178	[.069]
Other division [yes=1]	.065	.000	[.040]
Panel B: Workplace Climate and Ideas			
Aggregate Climate Index	-.106	.034	[.014]
	(.048)	(.029)	
Fostering New Ideas	-.068	-.021	[.509]
	(.059)	(.040)	

Notes: In Panel A, we report division-level aggregates as measured from the endline civil servants survey, and other division-level characteristics. In Panel B we report the aggregate workplace climate z-score and the sub-index z-score for fostering new ideas based on baseline data. In Column 3, the p-values are calculated from an underlying regression where standard errors are clustered at the division-organization level.

Table 4: Workplace Climate

OLS estimates, standard errors clustered by division-organization, p-values in brackets

A. Classroom Training (T1 vs C1)

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Classroom training (T1)	.180***	.189*	.191**	.194**	.156*	.113	.179*
	(.068)	(.108)	(.083)	(.092)	(.083)	(.169)	(.105)
	[.009]	[.083]	[.023]	[.037]	[.063]	[.505]	[.090]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in C1	-.049	-.066	-.038	-.047	-.037	.008	-.092
Observations (individual)	157	157	157	157	157	155	157

B. Division Training (T2 vs C2)

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Division Training (T2)	.003	.034	-.039	-.021	.036	.039	.030
	(.056)	(.075)	(.075)	(.075)	(.064)	(.112)	(.105)
	[.957]	[.656]	[.609]	[.780]	[.575]	[.727]	[.776]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in C2	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	363	363	363	363	363	359	363

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. In Panel A, the sample in all Columns is those bureaucrats that took part in the C1/T1 interventions. The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. Each specification controls for the gender of the respondent, and the randomization strata for the C1/T1 intervention. In Panel B, the sample in all Columns covers bureaucrats in C2 or T2 divisions. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division, whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table 5: Ideas Generation and Sharing

OLS estimates, standard errors clustered by team-organization, p-values in brackets

	T1 vs C1				T2 vs C2			
	(1) Freedom of expression... (index)	(2) Raising suggestions… (index)	(3) Discussing productivity improvements	(4) Individual officers originate ideas	(5) Freedom of expression... (index)	(6) Raising suggestions… (index)	(7) Discussing productivity improvements	(8) Individual officers originate ideas
Classroom training (T1)	.311**	.125	.136**	.120*				
	(.141)	(.113)	(.062)	(.062)				
	[.030]	[.274]	[.029]	[.055]				
Division Training (T2)					-.180	-.097	-.063	.053
					(.109)	(.099)	(.053)	(.040)
					[.101]	[.326]	[.238]	[.189]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C1)	3.43	4.69	.449	.211				
Control mean (C2)					3.55	4.68	.533	.202
Observations (individual)	118	118	275	269	268	268	624	616

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in Columns 1 to 4 covers all bureaucrats who attended the ten-day individual-level training (C1/T1) and are tracked to the endline civil servant survey. The samples in Columns 5 to 8 cover all bureaucrats who attended either the T2 division-level training or where assigned to a control division (C2). In Columns 1 and 5, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "In my work unit, I can freely express my thoughts," "In my work unit, expressing your true feelings is welcomed," "I proactively report coordination problems in the workplace to the management," and "I advise other colleagues against undesirable behaviors that would hamper job performance." In Columns 2 and 6, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "I raise suggestions to improve the unit's working procedure," "I make constructive suggestions to improve the unit's operation," "I proactively develop and make suggestions for issues that may influence the unit," "I proactively suggest new projects which are beneficial to the work unit," and "I proactively voice out constructive suggestions that help the unit reach its goals." In Columns 3 and 7, the dependent variable is the proportion of civil servants reporting to have monthly or more regular meetings to the question "Aside from training courses, how often would you say your division discusses ideas or ways to make work processes more efficient?". In Columns 4 and 8, the dependent variable is a dummy equal to one if the bureaucrat reports that individual officers commonly originate ideas. All columns report OLS estimates where standard errors are clustered at the division-organization level. In Columns 1 to 4 each specification controls for the gender of the respondent and the randomization strata for the C1/T1 intervention. In Columns 5 to 8 each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table 6: Administrative Processes and Task Completion

OLS estimates, standard errors clustered by division-organization

	(1) Quality of Procedures	(2) Adherance to Procedures	(3) Quality of Content	(4) Task Initiation	(5) Task Completion Rate	(6) Task Completion
At least one member of the division was assigned to classroom training T1	.286**	.322*	.133	-.006	.201	.110*
	(.144)	(.164)	(.178)	(.011)	(.149)	(.060)
	[.050]	[.052]	[.455]	[.588]	[.179]	[.071]
Division Training (T2)	.156	.132	.167	.002	-.174	-.039
	(.117)	(.170)	(.187)	(.015)	(.160)	(.071)
	[.187]	[.439]	[.375]	[.873]	[.280]	[.579]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean in other divisions	-.018	-.034	-.084	.987	3.48	.162
p-value: equality	[.445]	[.382]	[.875]	[.623]	[.110]	[.120]
Observations	286	286	286	627	627	627

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. In Columns 1 to 3 the unit of observation is the administrative file, where files are those that are open or in use after the ten-day standard training cycle. In Columns 4 to 6 the unit of observation is the task assigned to divisions in 2018 annual reports. Columns 1-3 (4-6) are inverse-weighted by the number of files (tasks) per division. All Columns report OLS estimates, where standard errors are clustered at the division-organization level. Columns 1-3 control for file characteristics (whether the assessor stated they had access to all the information needed to assess the file; a dummy stating that there were no challenges encountered in judging the quality of the file; the log number of files in the division; dummies for the file assessors; and the day on which the assessment was done). Column 4-6 includes controls for task characteristics (a dummy stating whether the task is a one-off task or a periodic/regular task; a dummy stating whether the reported task is a single task, or a bundle of tasks; a dummy stating whether the task requires coordination with stakeholders outside the organization; and the log number of tasks in the division). All specifications control for the number of bureaucrats from the division that attended the C1/T1 training.

Table 7: Quality of Action Plans and Innovation Proposed

Sample: Individual and division action plans
Panels report proportions, p-values in brackets

	New Training (T1)	Division Training (T2)	p-value [T1=T2]
Number of action plans	133	89	
Number of organizations represented	34	31	
Number of divisions represented	90	89	
Panel A: Quality dimensions (share of scores of 4 or 5, on a 1-5 scale)			
Lessons from training	.354	.154	[.060]
Level of detail	.367	.483	[.089]
Overall feasibility	.360	.292	[.303]
Degree of ambition	.289	.270	[.756]
Success fully defined	.252	.169	[.173]
Panel B: Focus of action plan (choose one)			
Routine work processes	.441	.326	[.089]
Data, monitoring, and oversight	.260	.170	[.113]
Communications inside and outside organization	.087	.045	[.238]
Financial management, procurement, resources	.063	.124	[.123]
Personnel management	.079	.134	[.181]
Other administrative	.071	.202	[.004]
Panel C: Nature of proposed innovation (choose one)			
Implementing existing practice	.391	.461	[.306]
Expanding scope of existing practice	.281	.292	[.862]
New innovation	.148	.180	[.539]
Panel D: Main implementation activities envisioned (choose all that apply)			
Meetings within the division	.459	.337	[.072]
Meetings with other divisions in the organization	.376	.348	[.677]
Acquire resources, dedicate time	.271	.382	[.081]
Provide training	.195	.303	[.065]
Meetings with actors outside the organization	.068	.157	[.032]
Individual behavior change	.060	.101	[.263]
Gather evidence or new ideas	.053	.079	[.437]

Notes: We report descriptives on action plans collected from trainees during the various trainings delivered during the study period. Individual-level action plans (T1) are restricted to the sample of individuals of professional grade, consistent with the grades of officers that are included in our baseline and endline surveys. In Panel A, for 'Success fully defined', proportion is the percentage of action plans coded as having success 'Fully defined' rather than 'Partially defined' or 'Not defined'. In Panels B and C, the descriptives refer to the main topic on which the action plan focused (B), and the nature of the proposed change (C). Coders were asked to select the option that was most applicable; figures may not sum to 1 due to rounding or the omission of the residual 'other' category. In Panel D the descriptives refer to the main activities envisioned by the action plan writer(s) to implement the action plan. For brevity, some smaller categories were aggregated into the categories presented here. P-values reported are for two-sided t-tests.

Table 8: Implementation of Action Plans

Sample (Panel A): Individual and division action plans
Sample (Panels B and C): endline civil servant survey
Panels report proportions, p-values in brackets

	New Training (T1)	Division Training (T2)	p-value [T1=T2]
Panel A: Ex ante obstacles envisioned (choose all that apply)			
Cooperation within the division	.466	.449	[.807]
Cooperation from other divisions in the org	.346	.393	[.475]
Resources, time, other	.331	.416	[.200]
Cooperation from actors outside the org	.098	.124	[.545]
Individual skills	.098	.112	[.728]
Panel B: Ex post obstacles to applying learning (choose all that apply)			
Division as a whole not supportive of implementing new ideas/practices	.125	.240	[.034]
Lack of resources/logistical constraints	.138	.287	[.034]
Putting training material into practice was too difficult	.504	.344	[.073]
Manager not supportive of implementing new ideas/practices	.567	.618	[.561]
Training material was not relevant to my work processes	.107	.141	[.477]
Panel C: Steps taken to implement action plan (choose all that apply)			
Discussed with superiors (e.g. Director, Chief Director)	.582	.723	[.188]
Discussed with other colleagues	.400	.518	[.200]
Worked as an individual to implement Action Plan	.185	.311	[.094]
Set up a team/committee	.174	.128	[.468]
Undertook additional feasibility research	.061	.010	[.150]

Notes: In Panel A we report descriptives on action plans collected from trainees during the various trainings delivered during the study period. Individual-level action plans (T1) are restricted to the sample of individuals of professional grade, consistent with the grades of officers that are included in our baseline and endline surveys. These relate to obstacles to success envisioned by the action plan writer(s). Coders were asked to select all options that were applicable. For brevity, some smaller categories were aggregated into the categories presented here. P-values reported in Panel A are for two-sided t-tests. Panels B and C report outcomes recorded from the endline survey of bureaucrats. In these panels when testing for differences between T1 and T2 we control for gender, education (having an undergraduate degree or not), whether the individual works for a human reseource managment division or other division, and randomization strata. Figures may not sum to one due to rounding or the omission of the residual 'other' category. P-values reported are for two-sided t-tests.

Figure 1: Study Timeline

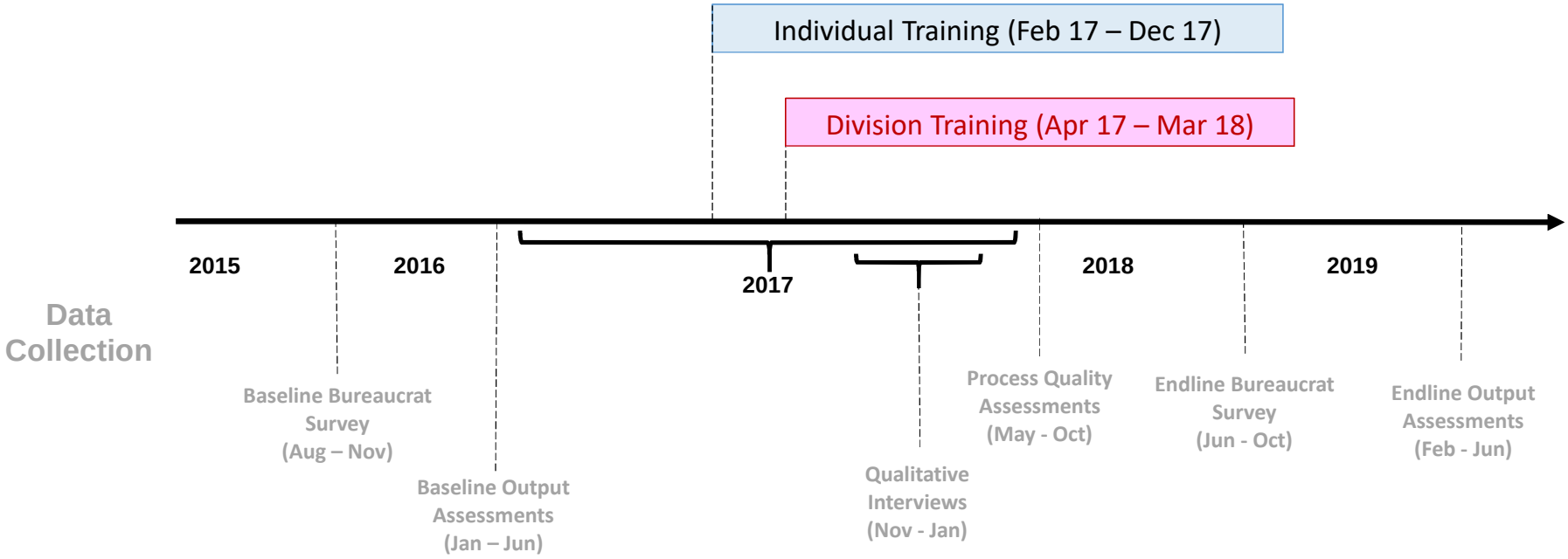
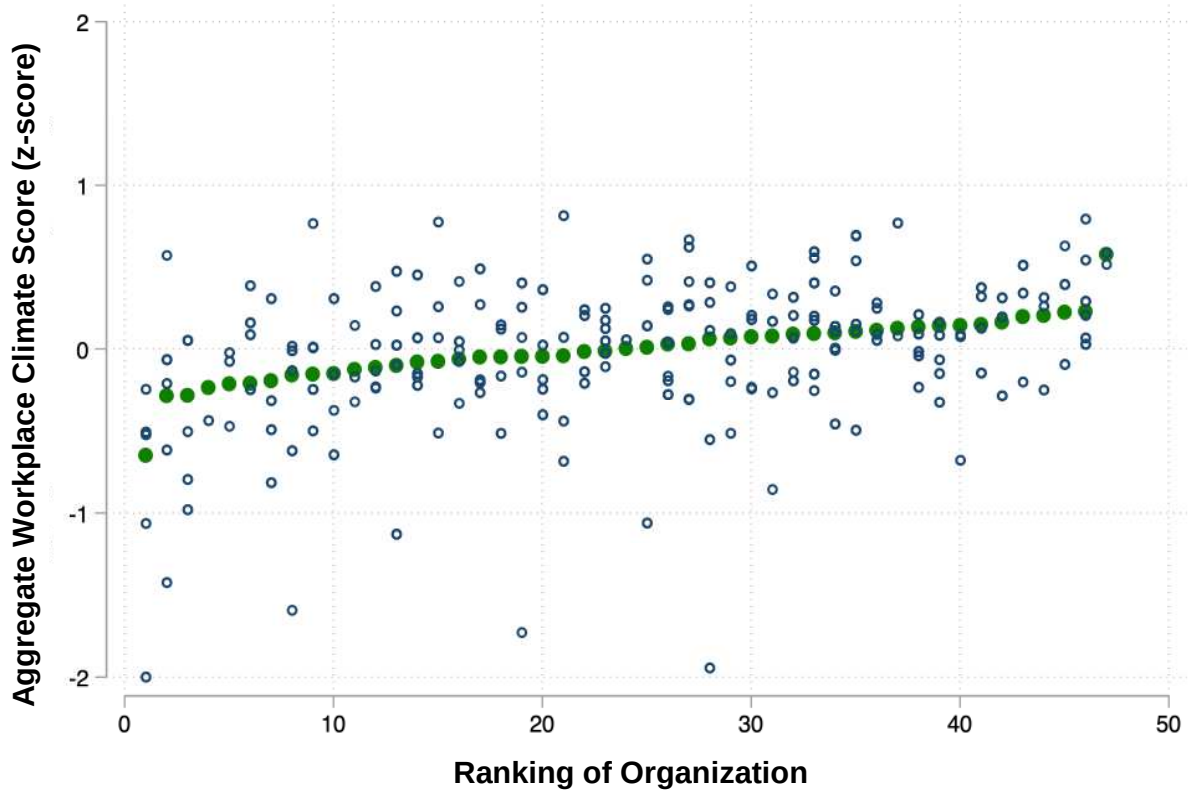
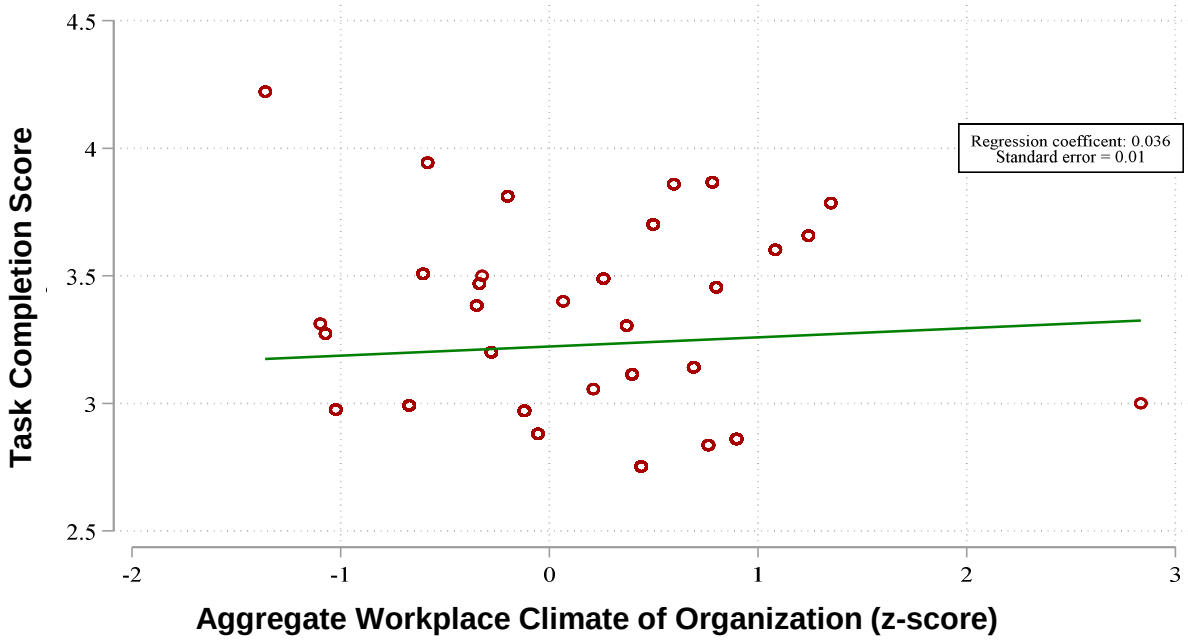


Figure 2: Workplace Climate of Bureaucracies

Panel A: Aggregate Climate Score Across Organizations and Divisions, Baseline

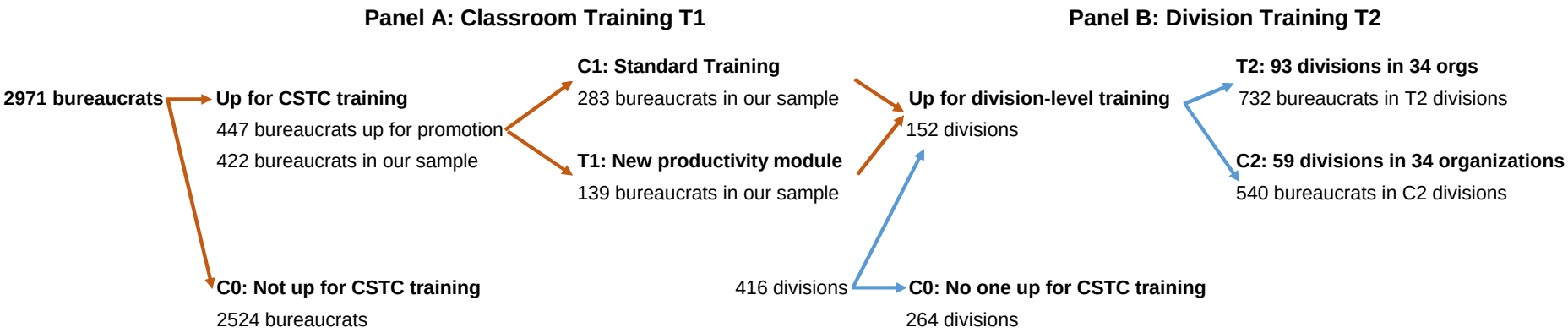


Panel B: Climate and Task Completion, Baseline



Notes: Panel A shows the baseline aggregate z-scores on the aggregate index of workplace climate. The solid dots are the index averages for each organization. The hollow dots are the index average for each division within the organization. Panel B presents a scatterplot of the baseline organization-level realizations of the aggregate index of workplace climate against the baseline organization-average task completion score. We show the line of best fit, the corresponding unconditional regression coefficient and its associated standard error.

Figure 3: Experimental Design



Panel C: Training Combinations of the 447 Bureaucrats up for Promotion

C1 + C2: 189 bureaucrats (173 sampled)

C1 + T2: 112 bureaucrats (110 sampled)

T1 + C2: 83 bureaucrats (77 sampled)

T1 + T2: 63 bureaucrats (62 sampled)

Table A1: Ghanaian Civil Service Organizations

Civil Service Organization	Sector	Annual Budget (US$)	Number of Divisions	Study Sample
Ministry of Education	Social Welfare	1,816,997,706	10	Yes
Ministry of Health	Social Welfare	866,470,461	13	Yes
Ministry of Interior	Security	316,952,632	9	Yes
Ministry of Roads and Highways	Infrastructure	290,864,351	13	Yes
Ministry of Energy and Petroleum	Infrastructure	250,093,305	12	Yes
Ministry of Power	Infrastructure	244,776,921	10	Yes
Ministry of Defence	Security	240,338,988	5	Yes
Ministry of Finance	Financial	191,879,079	18	Yes
Ministry of Water Resources, Works, and Housing	Infrastructure	144,544,988	10	Yes
Ministry of Food and Agriculture	Agriculture	116,361,440	19	Yes
Ministry of Transport	Infrastructure	111,944,055	10	Yes
Ministry of Local Government and Rural Development	Agriculture	88,944,422	11	Yes
Ministry of Communications	Administration	88,769,618	8	Yes
Ministry of Lands and Natural Resources	Environment	77,615,054	10	Yes
Ministry of Foreign Affairs	Administration	72,232,710	22	Yes
Ministry of Environment, Science, Technology, and Innovation	Environment	66,496,646	9	Yes
Ministry of Trade and Industry	Administration	53,132,044	13	Yes
Fisheries Commission	Agriculture	21,614,319	7	Yes
Ministry of Justice and Attorney-General's Department	Justice	19,678,444	10	Yes
Environmental Protection Agency	Environment	16,626,780	28	No
Ministry of Gender, Children and Social Protection	Social Welfare	12,307,746	7	Yes
Ministry of Employment and Labour Relations	Employment	11,554,163	10	Yes
Ministry of Youth and Sports	Social Welfare	9,269,598	5	Yes
Ministry of Tourism, Culture, and Creative Arts	Social Welfare	8,890,237	11	Yes
Ministry of Chieftaincy and Traditional Affairs	Administration	5,128,821	7	Yes
Office of the Head of Civil Service	Administration	2,703,970	8	Yes
Registrar-General's Department	Administration	2,465,892	18	Yes
Controller and Accountant-General's Department	Financial	1,828,623	13	No
Births and Deaths Department	Administration	1,392,568	11	No
Office of the President	Administration	543,531	10	No
Department of Feeder Roads	Agriculture	132,835	16	Yes
Public Works Department	Infrastructure	11,026	15	Yes
Department of Urban Roads	Infrastructure	8,599	10	Yes
Department of Children	Social Welfare	8,446	8	Yes
Public Records and Archives Admin Dept	Administration	6,490	8	Yes
Department of Cooperatives	Agriculture	.	5	Yes
Department of Factories Inspectorate	Administration	.	3	No
Department of Parks and Gardens	Environment	.	7	No
Department of Social Welfare	Social Welfare	.	14	Yes
Department of Women	Social Welfare	.	7	Yes
Geological Survey Department	Environment	.	12	No
Information Services Department	Administration	.	24	Yes
Labour Department	Employment	.	11	Yes
Public Sector Reform Secretariat	Administration	.	5	Yes
Town and Country Planning Department	Administration	.	8	Yes

Notes: A number of organization names changed across the course of the study, and so the names displayed represent administrative entities that may have come under distinct titles. Within these changes, there were a small number of combinations and separations of functions. Sector classifications are author defined. The budget figures are averages for 2015 to 2017. They are in US Dollars exchanged at a rate of US$1: Ghanaian Shilling 3.88, an average for the study period. For some organizations separate budget figures are not published and therefore not available. Division numbers come from baseline survey data for 2015. The Column Study Sample indicates whether the organization appears in our core specifications.

Table A2: Attrition

Dependent variable: Individual is retained in sample between training and endline
Robust standard errors in parentheses

	(1) Unconditional	(2) Controls	(3) Interactions	(4) Unconditional	(5) Controls	(6) Interactions
Classroom training (T1)	-.004	-.020	.128			
	(.038)	(.046)	(.186)			
Team Training (T2)				.025	.023	-.021
				(.025)	(.032)	(.116)
Gender [female == 1]		-.035	-.049		.005	-.015
		(.041)	(.049)		(.022)	(.048)
Working in Finance and Administration		-.018	.017		-.030	-.060
		(.060)	(.062)		(.031)	(.054)
Working in Policy, Planning, Monitoring and Evaluation		-.032	-.044		-.034	-.017
		(.076)	(.091)		(.037)	(.073)
Working in Research and Statistics		.068	.035		.012	-.078
		(.065)	(.080)		(.036)	(.085)
Working in Human Resources		.039	.011		.003	-.131
		(.076)	(.103)		(.039)	(.128)
Working in Other Divisions		.123	.044		.107***	-
		(.084)	(.063)		(.034)	-
Attrition rate		13.3			11.6	
Randomization strata for T1	Yes	Yes	Yes	No	No	No
Other controls	No	Yes	Yes	No	Yes	Yes
Joint p-value on interactions	-	-	[.335]	-	-	[.902]
Observations	422	322	322	944	944	944

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The outcome is whether the individual bureaucrat attrits between the training and the endline civil servant survey. Columns 1 to 3 focus on attrition between T1 and C1, and Columns 4 to 6 examine attrition of those in T2 relative to those in C2. For each we examine whether attrition is: (i) unconditionally correlated to treatment; (ii) is correlated to treatment conditional on bureaucrat characteristics; (iii) heterogeneous across individuals depending on these characteristics. In Columns 1 to 3 we control for the randomization strata for the C1/T1 intervention. In Columns 2, 3, 5 and 6 we control for whether an individual took a training entry exam, entry exam scores, and the type of exam taken. At the foot of Columns 3 and 6 we report the p-value on a test of the joint significance of all interactions with treatment dummies for T1 and T2.

Table A3: Feedback on Training

Sample: Classroom and division training assessments

	Standard clasroom training (C1)	New classroom training (T1)	Division training (T2)
Number of training sessions	17	11	85
Number of trainees	283	139	694
Number of organizations represented	40	36	34
Number of divisions represented	132	90	93
Training characteristics			
level of engagement of students	0.60	0.73	0.86
use of team work	0.00	0.82	0.80
use of role play	0.00	0.27	0.36
extent to which training was practical	0.40	0.73	0.78
use of civil service related examples	0.44	0.64	0.72
use of task-specific language	0.11	0.18	0.73
trainee questions implied a lack of basic understanding	0.00	0.18	0.00
quality of trainer's responses to questions	0.89	0.91	0.81

Notes: We report descriptives on the various trainings delivered during the study. The descriptives refer to the standard (C1) or new (T1) training module sessions, and division-level training (T2), subjectively evaluated by a training facilitator who observed the training. The descriptives report the proportion of sessions that scored 3.5 or higher on a five point Likert scale for each question.

Table A4: Workplace Climate and Bureaucrat Selection

OLS estimates, standard errors clustered by division-organization, p-values in brackets

A. Classroom Training, Tracked Panel (T1 vs C1)

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Classroom training (T1)	.173	.024	.188	.386**	.216	-.177	.414*
	(.113)	(.191)	(.147)	(.153)	(.146)	(.233)	(.222)
	[.131]	[.899]	[.207]	[.015]	[.147]	[.451]	[.069]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for baseline outcome	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in C1	-.049	-.066	-.038	-.047	-.037	.008	-.092
Observations (individual)	60	60	60	60	60	59	60

B. Division Training, Attendees (T2 vs C2)

Division Training (T2)	.027	-.001	-.013	.045	.049	-.100	.190*
	(.051)	(.078)	(.073)	(.078)	(.066)	(.116)	(.113)
	[.604]	[.987]	[.854]	[.560]	[.453]	[.388]	[.093]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in C2	.023	-.002	.076	.019	.017	.081	-.095
Observations (individual)	363	363	363	363	363	359	363

C. Division Training, Tracked Panel (T2 vs C2)

Division Training (T2)	-.027	-.035	-.076	.074	-.024	.188	.094
	(.093)	(.146)	(.119)	(.123)	(.110)	(.183)	(.189)
	[.774]	[.814]	[.524]	[.547]	[.825]	[.308]	[.619]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Contraolling for baseline outcome	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in C2	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	121	121	121	121	121	119	121

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. In Panel A, the sample in all Columns is those bureaucrats that took part in the C1/T1 interventions and are matched between the baseline and endline civil servant surveys. The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. Each specification controls for the gender of the respondent, and the randomization strata for the C1/T1 intervention. In Panel B, the sample in all Columns covers bureaucrats in C2 and T2 divisions, but the T2 treatment dummy is equal to one for those that actually took part in T2 division-level training. Each specification controls for the baseline value of the outcome, the gender of the respondent, the randomization strata for the C1/T1 intervention, whether the individual is treated with T1, the number of bureaucrats in the division, the share of women in the division, the share of bureaucrats in the division with an undergraduate degree, and whether the division is a Human Resources Division or an Other division. In Panel C, the sample in all Columns covers bureaucrats in C2 or T2 divisions. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the baseline measure of the outcome, the number of bureaucrats in the division, whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A5: Ideas Generation and Sharing, Components

OLS estimates, standard errors clustered by team-organization, p-values in brackets

A. Classroom Training (T1 vs C1)

	Freedom of expression:				Raising suggestions on:				
	(1) Express Thoughts	(2) Express Feelings	(3) Report on Coordination Issues	(4) Criticise Constructively	(5) Procedures	(6) Operations	(7) General Issues	(8) New Projects	(9) Goals
Classroom training (T1)	.046	.202***	-.000	.058	.062**	.028	.050*	.020	-.011
	(.043)	(.058)	(.081)	(.058)	(.030)	(.020)	(.026)	(.074)	(.027)
	[.291]	[.001]	[.996]	[.323]	[.046]	[.171]	[.059]	[.782]	[.673]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C1)	.938	.775	.835	.888	.938	.975	.950	.872	.975
Observations (individual)	118	118	117	118	117	118	118	115	118

B. Division Training (T2 vs C2)

Division Training (T2)	-.042	-.017	-.022	-.068	-.046**	-.025	-.012	-.009	-.005
	(.026)	(.048)	(.056)	(.046)	(.022)	(.017)	(.029)	(.049)	(.033)
	[.106]	[.723]	[.694]	[.138]	[.036]	[.136]	[.668]	[.860]	[.873]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	.967	.826	.837	.924	.978	.978	.946	.878	.935
Observations (individual)	268	267	265	266	267	268	265	261	267

Notes:*** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in Panel A covers all bureaucrats who attended the ten-day individual-level training (C1/T1) and are tracked to the endline civil servant survey. The sample in Panel B covers bureaucrats in C2 and T2 divisions. In Columns 1-9, the dependent variable is a dummy equal to one if the civil servant agrees/strongly agrees with the following questions, respectively: "In my work unit, I can freely express my thoughts," "In my work unit, expressing your true feelings is welcomed," "I proactively report coordination problems in the workplace to the management," "I advise other colleagues against undesirable behaviors that would hamper job performance," "I raise suggestions to improve the unit's working procedure," "I make constructive suggestions to improve the unit's operation," "I proactively develop and make suggestions for issues that may influence the unit," "I proactively suggest new projects which are beneficial to the work unit," and "I proactively voice out constructive suggestions that help the unit reach its goals." In Panel A, each specification controls for the gender of the participant, and the randomization strata for the C1/T1 intervention. In Panel B, each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A6: Other Outcomes

OLS estimates, standard errors clustered by team-organization, p-values in brackets

A. Classroom Training (T1 vs C1)

	Management Practices				Trust and Corruption			Job Satisfaction	
	(1) Aggregate management score	(2) Performance score	(3) Autonomy score	(4) Other practices score	(5) Trust Colleagues a Lot	(6) % of Projects Under Pressure to Divert Funds	(7) % of Projects in Which Observe Others Breaking Rules for Own Benefit	(8) Wants to Change Job	(9) Prefer Working in Public Sector
Classroom training (T1)	.112*	.188**	.092	.055	.055	-.632	-5.90	-.019	-.073
	(.067)	(.081)	(.073)	(.070)	(.043)	(1.32)	(4.30)	(.067)	(.061)
	[.097]	[.022]	[.208]	[.435]	[.209]	[.632]	[.172]	[.782]	[.238]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C1)	-.026	-.026	-.100	.049	.144	2.31	29.3	0.29	.413
Observations (individual)	268	268	268	268	270	176	221	264	271

B. Division Training (T2 vs C2)

Division Training (T2)	-.001	-.009	-.026	.021	-.029	.318	-3.44	-.049	.024
	(.079)	(.089)	(.098)	(.078)	(.035)	(.883)	(3.29)	(.042)	(.046)
	[.989]	[.920]	[.789]	[.789]	[.409]	[.719]	[.296]	[.250]	[.596]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	.018	.050	-.042	.046	.191	1.36	27.9	.263	.357
Observations (individual)	617	617	617	617	623	407	489	601	624

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in Panel A covers all bureaucrats who attended the ten-day individual-level training (C1/T1) and are tracked to the endline civil servant survey. The sample in Panel B covers bureaucrats in C2 and T2 divisions. In Columns 1-4, the dependent variable in each Column refers to an element of the management index z-scores, as defined in Appendix C Tables. In Column 1, the dependent variable corresponds to the aggregate management index. In Column 5, the dependent variable is equal to one if the respondent reports trusting their division colleagues 'a lot'. In Columns 6 and 7 the dependent variables are the proportion of projects respondents report they "were put under pressure to divert some of the funds," and "observed others breaking service rules for their own benefit," respectively. In Column 8, the dependent variable is a dummy variable equal to one if the respondent answers yes to the question "In the next two years, would you want to change your job?". In Column 9, the dependent variable is a dummy equal to one if the respondent agrees/strongly agrees with the statement, "Working in the public sector is generally better than working in the private sector." In Panel A, each specification controls for the gender of the participant, and the randomization strata for the C1/T1 intervention. In Panel B, each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A7: Organization of Work in Divisions

OLS estimates, standard errors clustered by team-organization, p-values in brackets

A. Classroom Training (T1 vs C1)

	(1) % of Tasks Assigned Directly By Seniors	(2) % of Tasks Working With Division Members	(3) Time at Work (hrs/wk)
Classroom training (T1)	3.95	-3.31	1.07
	(3.29)	(4.85)	(1.52)
	[.233]	[.497]	[.484]
Controls	Yes	Yes	Yes
Control mean (C1)	70.4	63.5	42.0
Observations (individual)	155	118	118

B. Division Training (T2 vs C2)

Division Training (T2)	1.95	-7.58**	-.509
	(2.47)	(3.42)	(.996)
	[.429]	[.029]	[.610]
Controls	Yes	Yes	Yes
Control mean (T2)	71.6	66.2	42.8
Observations (individual)	358	268	268

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in Panel A covers all bureaucrats who attended the ten-day individual-level training (C1/T1) and are tracked to the endline civil servant survey. The sample in Panel B covers all bureaucrats who attended either the T2 division training or where assigned to a control division (C2). In Column 1, the dependent variable is based on the question, “What percentage of your time is spent working on tasks given to you by your direct superior?” We asked respondents to provide an answer between 1 and 100. In Column 2 the dependent variable is based on the question, “What percentage of your time is spent working on tasks where you work closely with a team?” We asked respondents to provide an answer between 1 and 100. In Column 3 the dependent variable is based on the question, “What number of hours do you typically spend at the office in one week (including lunch and breaks)?”, requiring a numeric answer from respondents. In Panel A, each specification controls for the gender of the participant, and the randomization strata for the C1/T1 intervention. In Panel B, each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A8: Workplace Climate, Treatment Interactions

OLS estimates, standard errors clustered by team-organization

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Classroom training (T1)	.182*	.127	.214*	.176	.127	.551**	.168
	(.092)	(.154)	(.116)	(.135)	(.114)	(.260)	(.169)
Team Training (T2)	.017	.004	-.009	-.020	.058	.134	.064
	(.068)	(.084)	(.091)	(.090)	(.074)	(.123)	(.125)
T1 x T2	-.061	.121	-.145	-.024	-.033	-.597*	-.077
	(.114)	(.189)	(.144)	(.161)	(.155)	(.310)	(.205)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
T1 = T2 [p-value]	[.025]	[.362]	[.015]	[.085]	[.493]	[.083]	[.409]
Observations (individual)	363	363	363	363	363	359	363

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. Each specification controls for the gender of the respondent, and the randomization strata for the C1/T1 intervention. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division. At the foot of each Column we report the p-value on the equality of impacts of T1 and T2.

Table A9: Presence of Directors and Workplace Climate

OLS estimates, standard errors clustered by division-organization

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Division Training (T2)	-.007	.025	-.034	.020	-.039	.022	.015
	(.064)	(.089)	(.083)	(.093)	(.075)	(.145)	(.113)
Division Training (T2) x Director Present	.017	.016	-.007	-.072	.130**	.030	.026
	(.059)	(.083)	(.074)	(.090)	(.064)	(.147)	(.087)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean in C2	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	363	363	363	363	363	359	363

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. The sample in all Columns covers bureaucrats in C2 or T2 divisions. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the number of bureaucrats in the division, whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A10: Directors, Ideas Generation and Sharing

OLS estimates, standard errors clustered by division-organization

	(1) Freedom of expression... (index)	(2) Raising suggestions… (index)	(3) Discussing productivity improvements	(4) Individual officers originate ideas
Division Training (T2)	-.196	-.107	-.078	.057
	(.141)	(.150)	(.064)	(.052)
Division Training (T2) x Director Present	.024	.015	.024	-.008
	(.139)	(.169)	(.059)	(.055)
Controls	Yes	Yes	Yes	Yes
Control mean (C2)	3.55	4.69	.533	.202
Observations (individual)	268	268	624	616

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in each Column covers all bureaucrats who attended either the T2 division-level training or where assigned to a control division (C2). In Column 1, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "In my work unit, I can freely express my thoughts," "In my work unit, expressing your true feelings is welcomed," "I proactively report coordination problems in the workplace to the management," and "I advise other colleagues against undesirable behaviors that would hamper job performance." In Column 2, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "I raise suggestions to improve the unit's working procedure," "I make constructive suggestions to improve the unit's operation," "I proactively develop and make suggestions for issues that may influence the unit," "I proactively suggest new projects which are beneficial to the work unit," and "I proactively voice out constructive suggestions that help the unit reach its goals." In Column 3, the dependent variable is the proportion of civil servants reporting to have monthly or more regular meetings to the question "Aside from training courses, how often would you say your division discusses ideas or ways to make work processes more efficient?". In Column 4, the dependent variable is a dummy equal to one if the bureaucrat reports that individual officers commonly originate ideas. All Columns report OLS estimates where standard errors are clustered at the division-organization level. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the size of the division (number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A11: Directors, Administrative Processes and Task Completion

OLS estimates, standard errors clustered by division-organization

	(1) Quality of Procedures	(2) Adherance to Procedures	(3) Quality of Content	(4) Task Initiation	(5) Task Completion Rate	(6) Task Completion
Division Training (T2)	.115	.189	.336*	-.019	-.089	.004
	(.141)	(.196)	(.195)	(.094)	(.199)	(.018)
Division Training (T2)*Director Present	.098	-.075	-.293*	-.045	-.162	-.003
	(.173)	(.156)	(.159)	(.082)	(.197)	(.013)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean in other teams	-.048	-.107	-.140	.173	3.423	.960
Observations	286	286	286	627	627	627

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. In Columns 1 to 3 the unit of observation is the administrative file, where files are those that are open or in use after the ten-day standard training cycle. In Columns 4 and 5 the unit of observation is the task assigned to divisions in 2018 annual reports. Columns 1-3 (4-5) are inverse-weighted by the number of files (tasks) per division. All Columns report OLS estimates, where standard errors are clustered at the division-organization level. Columns 1-3 control for file characteristics (whether the assessor stated they had access to all the information needed to assess the file; a dummy stating that there were no challenges encountered in judging the quality of the file; the log number of files in the division; dummies for the file assessors; and the day on which the assessment was done). Column 4-5 includes controls for task characteristics (a dummy stating whether the task is a one-off task or a periodic/regular task; a dummy stating whether the reported task is a single task, or a bundle of tasks; a dummy stating whether the task requires coordination with stakeholders outside the organization; and the log number of tasks in the division). All specifications control for the number of bureaucrats from the division that attended the C1/T1 training.

Table A12: Action Plans, by Director Presence in T2

Sample (Panels A to E): Individual and division action plans
Sample (Panels F and G): endline civil servant survey
Panels report proportions, p-values in brackets

	T2 - Director Present	T2 - Director Not Present	p-value [equality of means]
Number of divisions	46	27	
Panel A: Quality dimensions (share of scores of 4 or 5, on a 1-5 scale)			
Lessons from training	.231	.111	[.489]
Level of detail	.435	.593	[.198]
Overall feasibility	.326	.333	[.950]
Degree of ambition	.304	.296	[.943]
Success fully defined	.200	.095	[.302]
Panel B: Focus of action plan (choose one)			
Routine work processes	.304	.259	[.686]
Data, monitoring, and oversight	.196	.185	[.914]
Financial management, procurement, resources	.065	.148	[.251]
Communications inside and outside organization	.043	.037	[.895]
Other administrative	.217	.222	[.962]
Personnel management	.174	.148	[.778]
Panel C: Nature of proposed innovation (choose one)			
Implementing existing practice	0.391	0.519	[.297]
Expanding scope of existing practice	0.326	0.296	[.795]
New innovation	0.217	0.111	[.258]
Panel D: Main implementation activities envisioned (choose all that apply)			
Meetings within the division	.304	.333	[.800]
Meetings with other divisions in the organization	.348	.370	[.849]
Acquire resources, dedicate time	.457	.222	[.046]
Provide training	.283	.370	[.443]
Meetings with actors outside the organization	.130	.185	[.534]
Individual behavior change	.109	.074	[.633]
Gather evidence or new ideas	.109	.037	[.288]
Panel E: Ex ante obstacles envisioned (choose all that apply)			
Cooperation within the division	.413	.444	[.800]
Cooperation from other divisions in the org	.370	.519	[.219]
Resources, time, other	.457	.370	[.479]
Cooperation from actors outside the org	.152	.074	[.334]
Individual skills	.109	.111	[.975]
Panel F: Ex post obstacles to applying learning (choose all that apply)			
Manager not supportive of implementing new ideas/practices	.070	.101	[.506]
Putting training material into practice was too difficult	.006	.013	[.701]
Lack of resources/logistical constraints	.196	.183	[.842]
Division as a whole not supportive of implementing new ideas/practices	.068	.038	[.561]
Training material was not relevant to my work processes	.042	.019	[.547]
Panel G: Steps taken to implement action plan (choose all that apply)			
Discussed with superiors (e.g. Director, Chief Director)	.156	.108	[.444]
Discussed with other colleagues	.187	.258	[.378]
Worked as an individual to implement Action Plan	.043	.033	[.872]
Set up a team/committee	.109	.027	[.098]
Undertook additional feasibility research	.006	.016	[.393]

Notes: We report descriptives on action plans collected from divisions during the various trainings delivered during the study period. In Panel A, for 'Success fully defined', proportion is the percentage of action plans coded as having success 'Fully defined' rather than 'Partially defined' or 'Not defined'. In Panels B and C, the descriptives refer to the main topic on which the action plan focused (B), and the nature of the proposed change (C). Coders were asked to select the option that was most applicable; figures may not sum to 1 due to rounding or the omission of the residual 'other' category. In Panel D the descriptives refer to the main activities envisioned by the action plan writer(s) to implement the action plan. For brevity, some smaller categories were aggregated into the categories presented here. In Panel E the descriptives relate to obstacles to success envisioned by the action plan writer(s). Coders were asked to select all options that were applicable. For brevity, some smaller categories were aggregated into the categories presented here. Panels F and G report outcomes from the endline survey of bureaucrats. In these panels when testing for differences between T1 and T2 we control for gender, education (having an undergraduate degree or not), whether the individual works for a human reseource managment division or other division, and randomization strata. Figures may not sum to 1 due to rounding or the omission of the residual 'other' category. P-values reported are for two-sided t-tests.

Table A13: T1 Spillovers and Workplace Climate

Sample: C0 bureaucrats

OLS estimates, standard errors clustered by division-organization

	(1) Culture Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Constant: C0 bureaucrats with no T1 division colleagues	.022	.039	.019	.040	.001	-.013	.035
	(.023)	(.029)	(.032)	(.029)	(.021)	(.045)	(.038)
At least one T1 trainee in division	-.010	.029	-.031	-.006	-.042	.179*	-.065
	(.067)	(.069)	(.091)	(.093)	(.060)	(.093)	(.084)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T2 (division)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations (individual)	1583	1583	1583	1583	1582	1564	1582

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. The sample in all Columns is those bureaucrats that were not up for promotion training (C0). The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. Each specification controls for the gender of participants, randomization strata for the C1/T1 intervention, and whether the division was assigned to T2.

Table A14: T1 Spillovers, Ideas Sharing and Generation

Sample: C0 bureaucrats

OLS estimates, standard errors clustered by division-organization

	(1) Freedom of expression... (index)	(2) Raising suggestions… (index)	(3) Discussing productivity improvements	(4) Individual officers originate ideas
Constant: C0 bureaucrats with no T1 division colleagues	3.52	4.58	.537	.063
	(.043)	(.059)	(.020)	(.041)
At least one T1 trainee in division	.079	.110	.016	.001
	(.100)	(.125)	(.702)	(.405)
Controls	Yes	Yes	Yes	Yes
Controlling for T2 (division)	Yes	Yes	Yes	Yes
Observations (individual)	1078	1078	2592	2565

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in all Columns is those bureaucrats that were not up for promotion training (C0). In Column 1, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "In my work unit, I can freely express my thoughts," "In my work unit, expressing your true feelings is welcomed," "I proactively report coordination problems in the workplace to the management," and "I advise other colleagues against undesirable behaviors that would hamper job performance." In Column 2, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "I raise suggestions to improve the unit's working procedure," "I make constructive suggestions to improve the unit's operation," "I proactively develop and make suggestions for issues that may influence the unit," "I proactively suggest new projects which are beneficial to the work unit," and "I proactively voice out constructive suggestions that help the unit reach its goals." In Column 3, the dependent variable is the proportion of civil servants reporting to have monthly or more regular meetings to the question "Aside from training courses, how often would you say your division discusses ideas or ways to make work processes more efficient?". In Column 4, the dependent variable is a dummy equal to one if the bureaucrat reports that individual officers commonly originate ideas. All Columns report OLS estimates where standard errors are clustered at the division-organization level. Each specification controls for the gender of participants, randomization strata for the C1/T1 intervention, and whether the division was assigned to T2.

Table A15A: Division Heterogeneity and Workplace Climate

OLS estimates, standard errors clustered by division-organization

Extent to Which Individuals Work on Different Processes

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Division Training (T2)	-.077	-.094	-.093	-.033	-.028	-.159	-.120
	(.121)	(.150)	(.162)	(.155)	(.122)	(.179)	(.181)
Division Heterogeneity	-.102	-.164	-.084	-.063	-.072	-.200**	-.108
	(.087)	(.145)	(.090)	(.094)	(.073)	(.086)	(.119)
Division Training (T2) x Division Heterogeneity	.079	.152	.038	-.009	.113	.341**	.017
	(.136)	(.180)	(.170)	(.152)	(.129)	(.161)	(.175)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	287	287	287	287	287	284	287

Trust in Division Colleagues

	(1) Climate Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Division Training (T2)	-.122	-.028	-.251	-.171	.020	.040	-.195
	(.115)	(.151)	(.153)	(.164)	(.130)	(.206)	(.162)
Division Heterogeneity	-.149	-.110	-.245	-.171	.044	-.149	-.270
	(.255)	(.392)	(.307)	(.357)	(.249)	(.356)	(.324)
Division Training (T2) x Division Heterogeneity	.461	.203	.809*	.594	.173	.156	.425
	(.332)	(.509)	(.419)	(.457)	(.330)	(.525)	(.409)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	286	286	286	286	286	283	286

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. The sample in all Columns covers bureaucrats in C2 or T2 divisions. The measures of heterogeneity are constructed as follows. The task-based measure of heterogeneity is constructed from individual responses to the question, Which of the following processes are you most closely involved with? They could select multiple types of processes (such as writing reports, answering correspondence, and responding to audit queries) from a showcard. From this we constructed a Herfindahl index, aggregated to the division level, on the share of bureaucrats working on different processes in the division at baseline. The trust-based measure of heterogeneity is based on question where we asked respondents, on a scale of 1 to 4, how much do you trust each of the officers in your division? We then created an index using the share of bureaucrats that reported as score of 4. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the size of the division (number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A15B: Division Heterogeneity and Workplace Climate

OLS estimates, standard errors clustered by division-organization

Acceptance of Low-level Corruption

	(1) Culture Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Division Training (T2)	-.110	-.120	-.080	-.276*	.066	.199	-.363**
	(.120)	(.150)	(.158)	(.157)	(.135)	(.180)	(.173)
Division Heterogeneity	-.312*	-.407	-.205	-.564**	-.035	-.115	-.562**
	(.165)	(.297)	(.198)	(.218)	(.253)	(.310)	(.281)
Division Training (T2) x Division Heterogeneity	.277	.416	.054	.690*	-.030	-.309	.755*
	(.238)	(.353)	(.318)	(.349)	(.318)	(.440)	(.387)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	287	287	287	287	287	284	287

Public Service Motivation

	(1) Culture Index	(2) Teamwork Climate	(3) Performance Climate	(4) Fostering New Ideas	(5) Perception of Management	(6) Stress Recognition	(7) Working Conditions
Division Training (T2)	-.021	-.000	-.073	-.023	.058	.051	-.103
	(.074)	(.088)	(.098)	(.100)	(.084)	(.130)	(.113)
Division Heterogeneity	.095	.099	.079	-.052	.183	.107	.209
	(.093)	(.136)	(.134)	(.129)	(.136)	(.143)	(.162)
Division Training (T2) x Division Heterogeneity	-.031	-.046	-.009	.249	-.174	-.312	-.125
	(.126)	(.183)	(.178)	(.166)	(.154)	(.226)	(.213)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controlling for T1 (individual)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	.033	-.038	.088	.063	.018	-.032	.017
Observations (individual)	281	281	281	281	281	278	281

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. All columns report OLS estimates where standard errors are clustered at the division-organization level. The dependent variable in each Column is either the z-score of the aggregate workplace climate index perceived by individuals (Column 1) or z-scores of the various sub-indices. The sample in all Columns covers bureaucrats in C2 or T2 divisions. The measures of heterogeneity are constructed as follows. On the extent to which individuals accept low-level corruption, we calculate the share of bureaucrats in the division that at baseline agree with the statement, 'The government regards low-level corruption as unfortunate but unavoidable until it can pay better salaries." On Public Service Motivation, we use the Perry [1996] sub-indices related to attraction to policy making, compassion for underprivileged people, commitment to the public interest, and self-sacrifice, to measure the average motivation of bureaucrats in the division. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the size of the division (number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A16: Division Heterogeneity, Ideas Sharing and Generation

OLS estimates, standard errors clustered by division-organization

	Extent to Which Individuals Work on Different Processes				Trust in Division Colleagues			
	(1) Freedom of expression... (index)	(2) Raising suggestions… (index)	(3) Discussing productivity improvements	(4) Individual officers originate ideas	(5) Freedom of expression... (index)	(6) Raising suggestions… (index)	(7) Discussing productivity improvements	(8) Individual officers originate ideas
Division Training (T2)	-.175	-.010	-.063	.060	-.492*	-.608*	-.050	.001
	(.281)	(.355)	(.091)	(.070)	(.271)	(.351)	(.085)	(.075)
Division Heterogeneity	.191	.247	.012	.028	-.635	-1.128	-.104	-.210
	(.222)	(.318)	(.066)	(.054)	(.654)	(.826)	(.180)	(.146)
Division Training (T2) x Division Heterogeneity	-.405	-.730*	.117	-.059	.066	.336	.250	.067
	(.274)	(.378)	(.094)	(.080)	(.796)	(1.01)	(.245)	(.215)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	1.77	2.33	.533	.202	1.77	2.33	.533	.202
Observations (individual)	496	496	490	487	495	495	489	487

	Acceptance of Low-level Corruption				Public Service Motivation			
	(1) Freedom of expression... (index)	(2) Raising suggestions… (index)	(3) Discussing productivity improvements	(4) Individual officers originate ideas	(5) Freedom of expression... (index)	(6) Raising suggestions… (index)	(7) Discussing productivity improvements	(8) Individual officers originate ideas
Division Training (T2)	.007	.186	.010	.077	-.543***	-.626**	.008	.000
	(.295)	(.374)	(.103)	(.087)	(.197)	(.248)	(.059)	(.046)
Division Heterogeneity	.700	1.184	.161	-.047	-.095	-.014	-.011	.058
	(.581)	(.774)	(.164)	(.145)	(.328)	(.395)	(.071)	(.056)
Division Training (T2) x Division Heterogeneity	-1.29*	-1.95**	.044	-.194	.132	.060	-.026	-.150**
	(.724)	(.949)	(.229)	(.205)	(.420)	(.530)	(.101)	(.070)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean (C2)	1.77	2.33	.533	.202	1.77	2.33	.533	.202
Observations (individual)	495	495	489	486	487	487	481	478

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. The unit of observation is the individual bureaucrat, where outcomes are recorded from the endline civil servant survey. The sample in each Column covers all bureaucrats who attended either the T2 division-level training or where assigned to a control division (C2). In Columns 1 and 5, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "In my work unit, I can freely express my thoughts," "In my work unit, expressing your true feelings is welcomed," "I proactively report coordination problems in the workplace to the management," and "I advise other colleagues against undesirable behaviors that would hamper job performance." In Columns 2 and 6, the dependent variable is the aggregated index of the proportion of civil servants agreeing and strongly agreeing with the following questions: "I raise suggestions to improve the unit's working procedure," "I make constructive suggestions to improve the unit's operation," "I proactively develop and make suggestions for issues that may influence the unit," "I proactively suggest new projects which are beneficial to the work unit," and "I proactively voice out constructive suggestions that help the unit reach its goals." In Columns 3 and 7, the dependent variable is the proportion of civil servants reporting to have monthly or more regular meetings to the question "Aside from training courses, how often would you say your division discusses ideas or ways to make work processes more efficient?". In Columns 4 and 8, the dependent variable is a dummy equal to one if the bureaucrat reports that individual officers commonly originate ideas. The measures of heterogeneity are constructed as follows. The task-based measure of heterogeneity is constructed from individual responses to the question, Which of the following processes are you most closely involved with? They could select multiple types of processes (such as writing reports, answering correspondence, and responding to audit queries) from a showcard. From this we constructed a Herfindahl index, aggregated to the division level, on the share of bureaucrats working on different processes in the division at baseline. The trust-based measure of heterogeneity is based on question where we asked respondents, on a scale of 1 to 4, how much do you trust each of the officers in your division? We then created an index using the share of bureaucrats that reported as score of 4. On the extent to which individuals accept low-level corruption, we calculate the share of bureaucrats in the division that at baseline agree with the statement, 'The government regards low-level corruption as unfortunate but unavoidable until it can pay better salaries." On Public Service Motivation, we use the Perry [1996] sub-indices related to attraction to policy making, compassion for underprivileged people, commitment to the public interest, and self-sacrifice, to measure the average motivation of bureaucrats in the division. All columns report OLS estimates where standard errors are clustered at the division-organization level. Each specification controls for the gender of the respondent, whether the individual is treated with T1, the size of the division (number of bureaucrats in the division), whether the individual has an undergraduate degree, and whether the individual works in Human Resources Division or an Other division.

Table A17A: Division Heterogeneity, Administrative Processes and Task Completion

OLS estimates, standard errors clustered by division-organization

	Extent to Which Individuals Work on Different Processes					
	(1) Quality of Procedures	(2) Adherance to Procedures	(3) Quality of Content	(4) Task Initiation	(5) Task Completion Rate	(6) Task Completion
Division Training (T2)	.009	-.361	-.116	.007	-.077	-.045
	(.259)	(.387)	(.364)	(.007)	(.310)	(.163)
Division Heterogeneity	-.438	-.548	-.399	-.000	-.062	-.092
	(.332)	(.487)	(.449)	(.005)	(.236)	(.142)
Division Training (T2) x Division Heterogeneity	.370	.812	.571	-.002	-.084	.018
	(.388)	(.542)	(.503)	(.006)	(.288)	(.155)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean in other teams	2.73	57.9	-.040	.960	3.42	.173
Observations	221	221	221	467	467	467

	Trust in Division Colleagues					
	(1) Quality of Procedures	(2) Adherance to Procedures	(3) Quality of Content	(4) Task Initiation	(5) Task Completion Rate	(6) Task Completion
Division Training (T2)	.215	.138	.461	.003	-.547**	-.139
	(.215)	(.324)	(.357)	(.005)	(.270)	(.118)
Division Heterogeneity	.436	.253	.917	.001	-.265	-.032
	(.451)	(.748)	(.710)	(.008)	(.288)	(.194)
Division Training (T2) x Division Heterogeneity	.061	.204	-.750	.009	1.70**	.416
	(.584)	(.801)	(.814)	(.011)	(.733)	(.320)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean in other teams	2.73	57.9	-0.04	.960	3.42	.173
Observations	221	221	221	467	467	467

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. In Columns 1 to 3 the unit of observation is the administrative file, where files are those that are open or in use after the ten-day standard training cycle. In Columns 4 and 5 the unit of observation is the task assigned to divisions in 2018 annual reports. Columns 1-3 (4-5) are inverse-weighted by the number of files (tasks) per division. All Columns report OLS estimates, where standard errors are clustered at the division-organization level. The measures of heterogeneity are constructed as follows. The task-based measure of heterogeneity is constructed from individual responses to the question, Which of the following processes are you most closely involved with? They could select multiple types of processes (such as writing reports, answering correspondence, and responding to audit queries) from a showcard. From this we constructed a Herfindahl index, aggregated to the division level, on the share of bureaucrats working on different processes in the division at baseline. The trust-based measure of heterogeneity is based on question where we asked respondents, on a scale of 1 to 4, how much do you trust each of the officers in your division? We then created an index using the share of bureaucrats that reported as score of 4. Columns 1-3 control for file characteristics (whether the assessor stated they had access to all the information needed to assess the file; a dummy stating that there were no challenges encountered in judging the quality of the file; the log number of files in the division; dummies for the file assessors; and the day on which the assessment was done). Column 4-5 includes controls for task characteristics (a dummy stating whether the task is a one-off task or a periodic/regular task; a dummy stating whether the reported task is a single task, or a bundle of tasks; a dummy stating whether the task requires coordination with stakeholders outside the organization; and the log number of tasks in the division). All specifications control for the number of bureaucrats from the division that attended the C1/T1 training.

Table A17B: Division Heterogeneity, Administrative Processes and Task Completion

OLS estimates, standard errors clustered by division-organization

Acceptance of Low-level Corruption

	(1) Quality of Procedures	(2) Adherance to Procedures	(3) Quality of Content	(4) Task Initiation	(5) Task Completion Rate	(6) Task Completion
Division Training (T2)	.190	.062	-.169	.005	-.523	-.199
	(.240)	(.344)	(.367)	(.006)	(.337)	(.149)
Division Heterogeneity	-.308	-.439	-.529	-.005	-.999**	-.495**
	(.460)	(.598)	(.400)	(.008)	(.408)	(.246)
Division Training (T2) x Division Heterogeneity	.002	.383	1.29*	.000	.907	.364
	(.633)	(.760)	(.759)	(.009)	(.724)	(.285)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean in other teams	2.73	57.9	-.040	.960	3.42	.173
Observations	221	221	221	455	455	455

Public Service Motivation

	(1) Quality of Procedures	(2) Adherance to Procedures	(3) Quality of Content	(4) Task Initiation	(5) Task Completion Rate	(6) Task Completion
Division Training (T2)	.175	.298	.270	.005	-.169	-.169
	(.169)	(.236)	(.259)	(.006)	(.185)	(.185)
Division Heterogeneity	.274	-.204	.374	.003	-.057	-.057
	(.351)	(.465)	(.443)	(.009)	(.390)	(.390)
Division Training (T2) x Division Heterogeneity	.035	.423	-.285	-.004	.201	.201
	(.379)	(.492)	(.436)	(.010)	(.441)	(.441)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean in other teams	2.73	57.9	-0.04	.960	3.42	.173
Observations	217	217	217	467	467	467

Notes: *** indicates significance at the 1% level, ** at the 5% level and * at the 10% level. In Columns 1 to 3 the unit of observation is the administrative file, where files are those that are open or in use after the ten-day standard training cycle. In Columns 4 and 5 the unit of observation is the task assigned to divisions in 2018 annual reports. Columns 1-3 (4-5) are inverse-weighted by the number of files (tasks) per division. All Columns report OLS estimates, where standard errors are clustered at the division-organization level. The measures of heterogeneity are constructed as follows. On the extent to which individuals accept low-level corruption, we calculate the share of bureaucrats in the division that at baseline agree with the statement, 'The government regards low-level corruption as unfortunate but unavoidable until it can pay better salaries." On Public Service Motivation, we use the Perry [1996] sub-indices related to attraction to policy making, compassion for underprivileged people, commitment to the public interest, and self-sacrifice, to measure the average motivation of bureaucrats in the division. Columns 1-3 control for file characteristics (whether the assessor stated they had access to all the information needed to assess the file; a dummy stating that there were no challenges encountered in judging the quality of the file; the log number of files in the division; dummies for the file assessors; and the day on which the assessment was done). Column 4-5 includes controls for task characteristics (a dummy stating whether the task is a one-off task or a periodic/regular task; a dummy stating whether the reported task is a single task, or a bundle of tasks; a dummy stating whether the task requires coordination with stakeholders outside the organization; and the log number of tasks in the division). All specifications control for the number of bureaucrats from the division that attended the C1/T1 training.

Table B1: Innovation in the Bureaucracy

Theme					
A. Hierarchy	*After expressing [an idea] they may say 'I'll call you later' and that just ends the idea there…So why force myself to try if it will be ignored? I usually get 'we will talk later' and when I go, [superior's name] postpones the time saying 'tomorrow'...[superior's name] then tells me that I am putting pressure… In the civil service before you take any action the powers can jump on you and you suffer that alone. Sometimes you go through transfers or they query you by giving you a note on why you are doing this and to explain. Things can then go on your file. (A2)*	*There is very limited space to do something, you will be seen as overshadowing the superior's portfolio or they may be someone who is not receptive to new ideas…Because of vertical reporting, ideas are stifled and it ends there… Rank within the service matters and from that who takes credit for an activity. So if the idea is proposed within a team or working group then the supervisor may not have interest. The interest may not be about the outcome but rather the process involved in getting to that outcome…This is all part of the institutional culture. (A4)*	*When I suggest anything that would enhance or facilitate progress of the job it is not taken. The status quo is to listen to the boss…I do not discuss with them; it is not appreciated so I do not even go there. (B1)*	*When you are the director they are the Gods and so for you to suggest things, they will think- who are you?! …They will think who are you? What have you seen? They may see it as rubbish. They will see it as you trying to outshine them. They are more comfortable talking about ideas from their level...Maybe in the future I can say something, when I am in a decision-making position. As one of my colleagues said 'for now we are just foot-soldiers who carry out the task. What you are told is what you do'… (D1)*	*Even if I think of ways, I do not see it materialise. Down the ladder, you have no voice. Informally you can recommend something to your director but formally I do not think I can change my directorate…The civil service is a robot - it is rigid…In the civil service, you voice certain things then you will be blacklisted. To play it safe, I have not tried. If you want to do things, then you have to be careful. (H5)*
B. Role of personal relations with superiors	*I sometimes take the ideas that juniors tell me and present it up and get things done for them… (C3)*	*It is easy to talk… I have never heard of being punished. If you have an idea, I believe in trying, how will you know? If you go to them, then you have to go in a nice way. Otherwise they could say "who are you to come to tell me and show me my work?"… This is not just about idea, but when you are brought closer to a project then they feel threatened to come on board and that you will take over. The only way they won't feel like this is if they trust that you won't take over. (I1)*	*I hear from colleagues that ideas are rejected and this lowers your morale. It all depends on if it is rejected in a nice way or is it done outright…This all depends on the superior you have and it boils down to the relationship you have with your superior. If you have a superior with a listening ear it is always good to think through things before you share it…This also all depends on the superior's character (A11)*		*It is quite easy to talk to them [about new ideas]…It all depends on the relationship, we are a very close unit and so we can voice things…In some organisations there is no friendliness, the juniors fear the superior. Even if a junior has an idea they may not tell if the boss is not friendly. If they are autocratic how can they attempt these things? (C4)*
C. Attitudes to innovation abd taking ideas from the ground up/horizontal conversations	*My former director, with some things I thought of I would say, and there were other things that I prefer to keep it at peer level [to be] free to discuss it. You would not be motivated to say anything as you know it would be killed as he had done this to other members. (A7)*	*You can bring ideas to the head but they may not take it on board…When I tried speaking the response was 'you are new'…Once we were having a team meeting and my colleague [name] suggested something and the head [superior's name] said 'I am the boss and you cannot decide'. In our informal little meetings, she told us she was not going to talk again at meetings and she has not done so since. Maybe [superior's name] saw this as a sign that we are not giving him respect or recognition as head…We only discuss the ideas among junior staff. We tried several times to talk with him and it is not working so why should we keep trying?...In our internal meetings we have ideas but we just discuss them among ourselves…There is fear of punishment and being victimised. We have heard that if you voice out or bring attention to things you are being monitored. Some are transferred to the village for speaking out. I know some people who have faced this and so it deters you from speaking out. (A13)*	*One aspect is the organisational culture…If you go to an institution where they do not listen to newly recruited staff, they do not have a voice. The civil service does not help in that manner as they make you feel that young people are not knowledgeable. It is not said, but in discussions, in the aura and ambience, you feel it. They may just say 'this is the way this is done' but you may realise that this is not right. I have personally suffered with this issue and I know other juniors who have suffered it. (A6)*	*You need to do your homework first and get the support from officers. The initial conversations about an idea start on the ground and then it is raised up…Having the support with colleagues really helps… (F2)*	*You do not have a voice, even if you can complain, they say you do not have the experience or know-how. These are flimsy excuses…You are then told to wait. Other colleagues tried to speak on how things were being done…So you then do not even ask the same question because you will get the same response. (A9)*
Some positive examples	*When there is a new idea, we discuss it and take it up. Seniors do encourage us during meetings. There is a channel in which you go through the director to the Chief Director. (D5)*	*I once had an idea about staff ID tags and I went to the technical advisor to the minister [rather than] the director for finance and administration. This is because I know the technical advisor can pop out the idea to the director of finance and administration and he was cool with me writing a proposal. (C4)*	*The communication we have is good with our director, she listens unlike other officers I have spoken to in other directorates and they are not so cool with information flow. (D1)*	*Our director is very active concerning these things…She asked me to compile ideas and to put in a recommendation. It means that I could then bring things out for her consideration…I share ideas with my colleagues such as [colleague's name]. You can also 'lead from behind'. You can push the idea. I do not need to wait for much or to get a pat on the back…Ideas will die in your mind without voicing them. (D2)*	*With my boss now, she is open to ideas, when it comes to the task, she is strict to meet it. She calls us and briefs us and asks for ideas that will make it fruitful. This all depends on the directorate and director, some want it to be like how it has been, so you cannot be innovative. It really depends... (E1)*

Table B2: Examples of Innovation

A. Simple Innovations	*Sometimes there are clashes in the meeting rooms as meetings are organised for the same times. No one took my idea seriously to have a noticeboard. I informed my committee and included the director as was instructed during the training. However, three years later it took hold when there was a massive mix up in the meeting schedule. I do not think anyone would remember that I had suggested it… (E2)*	*I enquired into the possibility of having anti-virus software as we are using our own laptops. I wrote the memo twice. I followed up with the IT guy about the software and then I learn that the memo has been changed to make it come from the director. This is not a matter that I am taking the credit for. It looks like there are interests that are being served. Do they think that I am not qualified to write a memo?...The whole system is messed up. This is about the sociology of the Ghanaian public sector… (A4)*	*During our tech meeting, we write minutes. The minutes' format was archaic, there was a different way to prepare the minutes, it was simpler and had clarity. I had the opportunity to write the minutes, [so I] use[d] that format and circulated it. I was then given the opportunity to share the format around and explain to staff how it works…Action speak louder than words. There are 7 technical staff members who now use this format. (A9)*	*About 12 years ago when I joined the bureau I had an idea about teaching officers about [topic]. I was willing to teach them during lunch hour breaks. I took it to the Minister's office. When they read the proposal I was asked 'do you not have enough work to do?' (F2)*	*I also came up with an idea to hold weekly meetings within my team, where we share our tasks and work on reports in a joint manner. Everyone will be assigned with a section to work on. (F1)*
B. System-wide Changes	*I have suggested to the director that the [multiple] agencies…we deal with tend to have more than one officer from [our ministry] in communication with them. Instead, we should have a way of collecting data and information from the agency with one officer [from our ministry]. After the director agreed to this, we had a meeting with the agency reps and it is now in action and took place a month ago. (D4)*	*We should have more engagement with [service beneficiaries] in the districts and I wish we would get up to date issues they are facing…up to date information on problems…By having more fieldtrips, we can see how to help them technically and enhance productivity… (A8)*	*Getting information from agencies [for monitoring and reporting] takes time…You need inputs from various agencies, this would [normally] be collected at the end [of the year], but we could compile information throughout the year and then compare things… (C1)*	*A strategy we have developed which is to meet with the agency people that will be doing the work, to make them understand what tasks they have. This has been beneficial as if it is a big report we are working on, then instead of waiting for it to arrive after one week and find that it is not how we want it to be, we will speak with them beforehand. I will then follow up with a call to see the progress…it has been very useful and something I have been doing (G2)*	*I do all the quarterly reports, regarding the agencies, their output is my input. They would delay in sending things to us, stating that they did not know of the deadlines, or that the letter arrived to them late. I suggested to my superior that we should send one letter at the beginning of the year outlining all the dates that we would require inputs from them. It made the agencies know they had things due…it helps us keep track of them and I could refer them to [the] letter. Also, sending the letter quarterly to agencies meant my boss would need to sign it, [as] sometimes my boss was not available and this would delay… (D2)*

Table B3: Feedback on Training

Standard Classroom Training (C1)	*Innovation is a pillar - coming up with new ideas that solves problems. This was something part of our [standard] training, because innovation is not welcomed much...things are withheld because 'we do not do this, we normally do it this way'. I learnt from the [standard] training that things need to be more innovative. (B2)*	*We were told to identify challenges hampering our work... The productivity module was theoretical… (C2)*	*[The] topic on productivity was very good, we had people who came to talk to us about the sector and share their ideas. The training was a good opportunity to bring people together to look at issues and look through certain technical things as well… (C3)*	*When I went on the [standard] training we were empowered to find resolutions to our problems…but the training was more lecturing, more interaction during the training is needed…We looked at ways to empower… (C6)*	*We learnt about efficient utilisation of resources, the wastage of resources…We learnt about work ethics, attendance, the level of respect between subordinates and superiors…The productivity module helped me realise, "garbage in and garbage out"…Productivity is the outcome; it is what matters. If it is negative it affects you, if it is positive it affects you… I learnt so much from the training…Every year we should go on the training including both seniors and juniors. (B1)*
New Classroom Training (T1)	*The training was not [too] theoretical…The training I had is still part of me...We need more practical training [like this]... (D2)*	*My way of doing things has changed...Now I have the time to take juniors through the work and let them know what I am expecting of them, rather than making the amendments and sending it to them with another officer. I talk the changes out with them... (G1)*			
Division-level Training (T2)	*We had a practical session to come up with ideas that did not need to rely on the Government of Ghana. There was an assignment that we had where we were tasked to develop something small that did not require money. … It was a useful as we are getting the ideas and then implementing them. Sometimes the ideas are there and there are a lot of ideas and people hold onto them even though they do not want to hold onto them. Therefore, after this training you can get the skills but the support is not there and so you just engage with other colleagues… (A2)*		*We learnt about productivity in the Civil Service and we watched videos of daily activities. It made me think about planning and prioritising my day and week. (C1)*	*We were just sent to a workshop on productivity and we then had the opportunity to present our ideas to the unit. We were told to set up a day by [our director] and then we had to tell them what went on during the day…[the] training was impactful, they played a video to us that showed different ministries. The director for RSIM in one ministry was saying how he shares work to his colleagues and decides on one person or another taking charge…When I saw this video and what was being done in RSIM, I was happy... (C4)*	*We had very useful trainings such as learning about productivity and gaps that we had. That training was very useful…They wanted us to look at the challenges we had such as how information is disseminated…We have information being kept at the top such as tasks and general information… One day was not enough; it was a rush. It would have taken at least 3 days. Even if it was 2 days that would have been fine… (A13)*

Table B4: Sentiments on the Presence of Directors in the Division-level Training (T2)

Negative Sentiment	*[Being] in the same room is not good. Better to keep it separate, have one for juniors, one for seniors and then all together in one room…If this had been a meeting situation, I would of not said anything. You are sitting with the big people, they know the job, they have the knowledge so who are you to talk, what have you done?...these are concerns to me, for there is fear of punishment… (A6)*	*The colleagues who you can talk to are not the same as the bosses…I have not voiced [my idea] to anyone that can implement it…When gathered as a group it is difficult to talk. When in an office it is easier to share, but all of us together, it does not happen… (C1)*	*With your superior it was not beneficial. You had to be cautious about the superior-subordinate relationship. I did not see the benefit. I could not voice my feelings…We had poor communication…we have tried but if superiors continue like this…The directorate is not for one person it is for all of us…[Our] action plans focused on having clear instructions, effective communication and team work. We had our superiors there as well. We could do it - everything is a 'can', but…only if the head of the directorate comes to our level. (G6)*	
Positive Sentiment	*I would prefer this type of training more than one with my colleagues but without my superior. The superior has different opinions and so without them you can talk so there is even a productivity issue and as the boss is not present he will throw the idea away...It should be enforced that they must be present because other leaders will shy away but the things not making the place a functional environment will be addressed…With the training, superiors should be involved from the beginning to the end stage. So any action plan should be the responsibility of the superior to make sure that the tasks are implemented… (A2)*	*[It] was okay, it was good that we were all together, we have junior officers and they get to know what is expected of us. (G1)*	*It is good to be in a training with your boss. He has more ideas and he made some comments that were helpful. The director was in my group. (A10)*	*It was also useful having our director present and so we found ourselves as a team. We were not frightened by his presence and he could put his ideas out there and share his views… (A13)*

Table C1: Organizational Culture and Comparison to Sexton et al. [2006]

Sub-index	Scale items used	Scale items used in Sexton et al. [2006]
Teamwork climate (1-5)	It is easy for personnel here to ask questions when there is something that they do not understand	It is easy for personnel in this ICU to ask questions when there is something that they do not understand.
	Disagreements in this division are resolved appropriately (i.e., not who is right, but what is best for the service)	Disagreements in this ICU are resolved appropriately (i.e., not who is right, but what is best for the patient)
	The managers and other officers here work together as a well-coordinated team	The physicians and nurses here work together as a well-coordinated team.
	Communication breakdowns that lead to delays in delivery of services are common	
		I have the support I need from other personnel to care for patients.
		Nurse input is well received in this ICU.
		In this ICU, it is difficult to speak up if I perceive a problem with patient care.
Performance climate (1-5)	The culture in this division makes it easy to learn from the errors of others	The culture in this ICU makes it easy to learn from the errors of others.
	You know the proper channels to direct questions regarding correct bureaucratic process in this division	I know the proper channels to direct questions regarding patient safety in this ICU.
	You receive appropriate feedback about your performance	I receive appropriate feedback about my performance.
	Bureaucratic errors are handled appropriately in this division	Medical errors are handled appropriately in this ICU.
	You are encouraged by your colleagues to report any work concerns you may have	I am encouraged by my colleagues to report any patient safety concerns I may have
	You would feel happy being served as a Ghanaian citizen by this division	I would feel safe being treated here as a patient.
		In this ICU, it is difficult to discuss errors.
Stress recognition (1-5)	Fatigue impairs your performance during high-pressure situations (e.g. when there are heavy demands on your division)	Fatigue impairs my performance during emergency situations (e.g., emergency resuscitation, seizure).
		I am more likely to make errors in tense or hostile situations.
		When my workload becomes excessive, my performance is impaired.
		I am less effective at work when fatigued.
Perceptions of management (1-5)	Staff (divisional management) doesn't knowingly compromise division services	Hospital management does not knowingly compromise the safety of patients.
	Staff (divisional management) supports your daily efforts	I am provided with adequate, timely information about events in the hospital that might affect my work.
	You get adequate, timely info about events that might affect your work from your division.	Hospital administration supports my daily efforts.
	Your performance has no influence on your career progression	The levels of staffing in this clinical area are sufficient to handle the number of patients
Working conditions (1-5)	All the necessary information for diagnostic and effective decision making is routinely available to you	All the necessary information for diagnostic and therapeutic decisions is routinely available to me.
	Trainees in your division are adequately supervised	Trainees in my discipline are adequately supervised.
		This hospital constructively deals with problem physicians and employees.
		This hospital does a good job of training new personnel
Fostering new ideas (1-5)	Your suggestions about work place productivity would be acted upon if you expressed them to management	
	Staff (divisional management) in this division are quick to adopt (are open to) new ways of doing things.	
	You can see lots of ways to make your division work better.	
Job satisfaction (1-5)		This hospital is a good place to work.
		Working in this hospital is like being part of a large family.
		I like my job.
		I am proud to work at this hospital.
		Moral in this ICU area is high.

Notes: Performance climate is referred to as the safety climate in Sexton et al. [2006]. Text in (parentheses) indicates item phrasing administered to non-management-level staff; all other items are identical for all respondents.

Table C2: Administrative Processes

Panel A: Quality of Procedure

Component	Questions for Assessment	Score 1	Score 3	Score 5
File Ladder	How complete is the file ladder? (Each transfer should be documented.)	0-19%	40-59%	80-100%
	Does each step in the file ladder have dates (each transfer is associated with a date)	0-19%	40-59%	80-100%
Folios	Are folios within the file organised and numbered consecutively?	0-19%	40-59%	80-100%
Memo and Minutes	Where applicable, are minutes, memos and other necessary records present and complete (including from whom, to whom and signature)?	0-19%	40-59%	80-100%
Correspondence	What proportion of incoming correspondence has an organizational stamp/date/signature?	0-19%	40-59%	80-100%
	What proportion of outgoing correspondence has a despatch stamp/date/signature?	0-19%	40-59%	80-100%

Panel B: Quality of Content

Component	Questions for Assessment	Score 1	Score 3	Score 5
Background to Issue	Background to issues	Very poor	Neither poor or good	Very good
Course of Action	Clearly outlining what courses of action are available or taken	Very poor	Neither poor or good	Very good
Logical Flow	The file is organised in a logical flow (where applicable, with an issue arising, being treated consecutively, and then resolved)?	Very poor	Neither poor or good	Very good
Choices	Choices are based on evidence in file	Very poor	Neither poor or good	Very good
Action Taken	Clarity on who should take actions at each stage	Very poor	Neither poor or good	Very good
Clear Deadline	What proportion of relevant materials have a clear deadline	Very poor	Neither poor or good	Very good

Table C3: Modules in Standard and New Training Curriculum

Module title	Indicative content/learning goals
Administrative writing	How to write handing-over notes, memos, letters
Policy development and analysis	Understanding policy problems; policy life cycle; policy development in Ghana; Role of civil servants in policy development
Administrative principles and instructions	Civil Service Code of Conduct; general principles of the Civil Service; rules for promotions, secondments, etc.
Human relations	Navigating the relationships required for working in the Civil Service; different styles of working and communication
Organizational security and safety	Classification and handling of records; confidentiality and disclosure of classified material
Team building	What is a team; essentials and features of teams; stages of development; personalities with a team; motivating a team
Work ethics and work standards	Principles of good work ethic; definition of work standards and how they fit into organizational practice; steps to develop work standards
Standard Productivity Module	Concept of productivity; definitions of *kaizen* and related concepts; link between productivity and national development; symptoms of low productivity; lessons learned from other African countries; stages of the productivity movement; examples from the Civil Service
New Productivity Module	Defining and measuring productivity; evidence on productivity variations within Ghana's Civil Service; methods for diagnosing potential work process improvements (e.g. 'fishbone' diagrams for identifying root causes of problems), with application by trainees to their teams; motivational video showing examples of bottom-up work process innovation from across the Civil Service; developing action plans and role playing.

Notes: Authors' synthesis of CSTC training materials. Note that the standard ten-day training curricula vary slightly in their module composition according to trainee grade and trainer availability.

Table C4: Standard and New Action Plan Templates

Standard action plan (C1)	New action plan (T1, T2)
1. What are some of your job responsibilities in your organisation?	*1. Inputs and Outputs* - 1.1 What are your inputs to your work? List the set of activities you undertake in your work day. - 1.2 What are your outputs in your work? What do you produce from these activities?
2. What challenge(s) do you encounter when you want to meet these responsibilities?	2. Problem Statement -2 .1 In what part of my division's work would I like to improve productivity?
3. How can you resolve these challenges?	3. Improving Productivity - 3.1 What needs to change to bring about this improvement? - 3.2 What are the main obstacles to making the necessary change(s)? - 3.3 What steps can I take to solve the problem? When? - 3.4 What help or support do I need from others? - 3.5 What new skills or knowledge, if any, will I need to acquire to carry out this initiative? - 3.6 How will I know when I have succeeded?
4. After this scheme of service course, when can you implement the above mentioned changes?	4. What are you going to do on Monday morning? - 4.1 What are you going to do on Monday morning to start your productivty improvement?

Notes: Authors' summary of action plan templates; both templates are two pages long in print form. All trainees are given a status quo action plan at the end of the 10-day training. The action plan template is identical for both treatment arms, but trainees in the division-based arm (T2) complete a single joint action plan for their entire division.

Table C5: Defining Management Practices

Management Practice	Topic	Indicative Question	Score 1	Score 3	Score 5
Autonomy/Discretion	**Roles**	Can most senior staff in your division make substantive contributions to the policy formulation and implementation process?	Senior staff do not have channels to make substantive contributions to organisational policies, nor to the management of their implementation.	Substantive contributions can be made in staff meetings by all senior staff but there are no individual channels for ideas to flow up the organisation.	It is integral to the organisation's culture that any member of senior staff can substantively contribute to the policies of the organisation or their implementation.
		When senior staff in your division are given tasks in their daily work, how much discretion do they have to carry out their assignments? Can you give me an example?	Officers in this division have no real independence to make decisions over how they carry out their daily assignments. Their activities are defined in detail by senior colleagues or organisational guidelines.	Officers in this division have some independence as to how they work, but strong guidance from senior colleagues, or from rules and regulations.	Officers in this division have a lot of independence as to how they go about their daily duties.
	Flexibility	Does your division make efforts to adjust to the specific needs and peculiarities of communities, clients, or other stakeholders?	The division uses the same procedures no matter what. In the face of specific needs or community/ client peculiarities, it does not try to develop a 'better fit' but automatically uses the default procedures.	The division makes steps towards responding to specific needs and peculiarities, but stumbles if the specific needs are complex. Often, tailoring of services is often unsuccessful.	The division always redefines its procedures to respond to the needs of communities/ clients. It does its best to serve each individual need as best as it can.
		How flexible would you say your division is in terms of responding to new and improved work practices?	There is no effort to incorporate new ideas or practices. When practice improvements do happen, there is no effort to disseminate them through the division.	New ideas or practices are sometimes adopted but in an ad hoc way. These are sometimes shared informally or in a limited way, but the division does not actively encourage this or monitor their adoption.	Seeking out and adopting improved work practices is an integral part of the division's work. Improvements are systematically disseminated throughout the division and their adoption is monitored.
Incentives/Monitoring	**Performance Incentives**	Given past experience, how would under-performance be tolerated in your division?	Poor performance is not addressed or is inconsistently addressed. Poor performers rarely suffer consequences or are removed from their positions.	Poor performance is addressed, but on an ad hoc basis. Use of intermediate interventions, such as training, is inconsistent. Poor performers are sometimes removed from their positions under conditions of repeated poor performance.	Repeated poor performance is systematically addressed, beginning with targeted intermediate interventions. Persistently poor performers are moved to less critical roles or out of the organisation.
		Given past experience, are members of [respondent's organisation] disciplined for breaking the rules of the civil service?	Breaking the rules of the civil service does not carry any consequences in this division. Guilty parties do not receive the stipulated punishment.	An officer may break the rules infrequently and not be punished. An officer who regularly breaks the rules may be disciplined, but there would be no other specific actions beyond this. The underlying drivers of the behaviour can persist indefinitely.	Any officer who breaks the rules of the civil service is punished; the underlying driver is identified and rectified. On-going efforts are made to ensure the issue does not arise again.
		Does your division use performance, targets, or indicators for tracking and rewarding (financially or non-financially) the performance of its officers?	Officers in the division are rewarded (or not rewarded) in the same way irrespective of their performance.	The evaluation system awards good performance in principle (financially or non-financially), but awards are not based on clear criteria/processes.	The evaluation system rewards individuals (financially or non-financially) based on performance. Rewards are given as a consequence of well-defined and monitored individual achievements.
	Monitoring	In what kind of ways does your division track how well it is delivering services? Can you give me an example?	Measures tracked are not appropriate or do not indicate directly if overall objectives are being met. Tracking is an ad hoc process and most processes aren't tracked at all. Tracking is dominated by the head of the division.	Performance indicators have been specified but may not be relevant to the division's objectives. The division has inclusive staff meetings where staff discuss how they are doing as division.	Performance is continuously tracked, both formally with key performance indicators and informally, using appropriate indicators and including many of the divisional staff.

Table C5 Continued: Defining Management Practices

Management Practice	Topic	Indicative Question	Score 1	Score 3	Score 5
Other	**Staffing**	Do you think about attracting talented people to your division and then doing your best to keep them? For example, by ensuring they are happy and engaged with their work.	Attracting, retaining and developing talent throughout the division is not a priority or is not possible given service rules.	Having top talent throughout the division is seen to be a key way to effectively deliver on the organisations mandate but there is no strategy to identify, attract or train such talent.	The division actively identifies and acts to attract talented people who will enrich the division. They then develop those individuals for the benefit of the division and try to retain their services.
		If two senior level staff joined your division five years ago and one was much better at their work than the other, would he/she be promoted through the service faster?	The division promotes people by tenure only, and thus performance does not play a role in promotion.	There is some scope for high performers to move up through the service faster than non-performers in this division, but the process is gradual and vulnerable to inefficiencies.	The division would certainly promote the high-performer faster, and would rapidly move them to a senior position to capitalise on their skills.
		Is the burden of achieving your division's targets evenly distributed across its different officers, or do some individuals consistently shoulder a greater burden than others?	A small minority of staff undertake the vast majority of substantive work within the division.	A majority of staff make valuable inputs, but it is by no means everyone who pulls their weight.	Each member of the division provides an equally valuable contribution, working where they can provide their highest value.
		Would you say that senior staff try to use the right staff for the right job?	Often tasks are not staffed by the appropriate staff. Staff are allocated to tasks either randomly, or for reasons that are not associated with productivity.	Most jobs have the right staff on them, but there are organisational constraints that limit the extent to which effective matching happens.	The right staff are always used for a task.
	Targeting	Does your division have a clear set of targets derived from the organization's goals and objectives? Are they used to determine your work schedule?	The division's targets are very loosely defined or not defined at all; if they exist, they are rarely used to determine our work schedule and our activities are based on ad hoc directives from senior management.	Targets are defined for the division and its individual officers (managers and staff). However, their use is relatively ad hoc and many of the division's activities do not relate to those targets.	Targets are defined for the division and individuals (managers and staff) and they provide a clear guide to the division and its staff as to what the division should do. They are frequently discussed and used to benchmark performance.
		When you arrive at work each day, do you and your colleagues know what their individual roles and responsibilities are in achieving the organisation's goals?	No. There is a general level of confusion as to what the organisation is trying to achieve on a daily basis and what individual's roles are towards those goals.	To some extent, or at least on some days. The organisation's main goals and individual's roles to achieve them are relatively clear, but it is sometimes difficult to see how current activities are moving us towards those.	Yes. It is always clear to the body of staff what the organisation is aiming to achieve with the days activities and what individual's roles and responsibilities are towards that.