CSAE Working Paper WPS/2021-08

Fostering Innovation through Empowered Workers - Experimental Evidence from the Bangladeshi Garment Industry

Vanessa Schreiber*

May 2021

Abstract

Informational constraints are often a crucial barrier to adoption of innovations in firms. How can firms overcome this barrier and increase upward flows of ideas? This paper, the first to examine this issue in the context of a manufacturing firm in a low-income country, assesses the impact on information flows from two managerial interventions that encouraged 1,600 line workers in a Bangladeshi garment factory to submit their ideas to improve operations and well-being at work. One intervention gave randomly selected workers time to discuss ideas with co-workers at free lunches; the other gave randomly selected supervisors communications training, and asked them to encourage their production line workers to develop ideas. I find that the lunch intervention fostered the upward flows of ideas, and led to improved worker well-being. By contrast, the supervisor-led meetings failed to accelerate the use of more efficient peer interactions; moreover, the presence of the supervisors undermined workers' critical voice about their bosses. The evidence suggests that the effectiveness of the lunches was driven by changes in workers' aspirations and the perceived ease of submission. By contrast, a perceived deterioration in the voice climate undercut the effectiveness of the supervision-led meetings. I also find that specific seating arrangements opposite experienced friends during lunches can be particularly beneficial for inexperienced workers.

Keywords: organizational learning, innovation, management, empowerment, Bangladesh.
JEL Classification: L2, M5, O15, D8.

*Vanessa Schreiber, University of Oxford, vanessa.schreiber@economics.ox.ac.uk. I am grateful to my supervisors, Christopher Woodruff and Simon Quinn, for their insightful feedback, advice and support. I also thank Raffaella Sadun, Sanjay Jain, Martin Williams, Meg Meyer, Lukas Hensel, Daniela Scur, Thomaz Teodorovicz, Florian Englmaier, Kathryn Shaw and seminar participants at the Empirical Management Conference, Cedefop, Eurofound and IZA Conference on Workplace and Management Practices, the Oxford's CSAE Workshop and Oxford's Firms, Labour and Organisations Research Group for their comments and suggestions. I am grateful to the field managers and enumerator team at the Innovations for Poverty (IPA) office in Bangladesh for their hard work on this project. I thank the the participating factory for their support and cooperation. I am thankful for generous financial support of this research from the Private Enterprise Development in Low-Income Countries and the ERC Advanced Grant RMGPP. Ethical approval was obtained from the Oxford Department of Economics' Departmental ch Ethics Committee (DREC) with reference ECONCIA19-20-24.

1 Introduction

"Workers on the production line might observe quality-defect problems that become apparent only on the shop floor, or notice a machine that is sometimes idle, or a surplus stock of raw materials that could be used." (Hayek, 1945)

As Hayek stated in 1945, knowledge valuable to an organization can come from people at all levels, including the lowest; indeed, valuable knowledge often surfaces from providing information about things that are seemingly trivial. Hayek's view reflects a problem that continues to hamper many organizations today. Information that would be helpful for the firm fails to rise up through the hierarchy. In this paper, I examine whether changing management practices to encourage bottom-up innovation leads to beneficial information flows from production workers on lower hierarchical levels to higher-level management with the power to implement ideas. I provide the first experimental evidence on impediments to idea-driven voice and change in manufacturing firms in low-income countries.

What limits communication of ideas from production workers to managers in hierarchical organizations and businesses? Workers may face costs of effort and time to develop ideas, arising from, for example, the inability to generate suggestions in an organizational language that managers can understand and reward (Weber and Camerer, 2003), or opportunity costs of time (Battiston and Kirchmaier, 2018). Workers may also face social costs that limit the diffusion of information, arising from, for example, reputational concerns (Bursztyn and Jensen, 2017), fear of signaling incompetence (Edmondson, 1999), or class differences between managers and workers, and altruism or loyalty toward in-group co-workers (Hjort, 2014). To address these issues, I designed experiments to assess the effectiveness of voice-enhancing, managerial interventions that target these two distinct cost categories – time and effort-related costs, and social costs – both separately and in combination. I set up a suggestion system on six production floors of a Bangladeshi garment factory that employs 1,588 workers on 51 production lines. The system offered financial awards for the top three ideas submitted each month, and displayed the winners' names to recognize their contributions. In this setting, I conducted a randomized

controlled trial (RCT) with two nested experiments between December 2019 and March 2020, using a cross-randomized design. The first experiment ("lunch treatment"), targeted costs of effort and time. Following the methodology of Sandvik et al. (2020), I randomly assigned workers into two groups. Workers in the treatment group were given a free lunch that was set up once per month to offer a dedicated time with co-workers to discuss suggestions for improvements. Assisted by external facilitators, workers in the treated group were asked to jointly discuss suggestions, and to fill out suggestion forms during the lunches. Workers in the control group also received a free take-away lunch (for fairness) and a suggestion form, but they were not provided with a fixed time to generate ideas or structured sessions involving external facilitators. The second experiment ("huddle treatment") targeted social costs by involving middle managers, who have been found in the literature to be an important part of the production function.[1] Supervisors on randomly selected production lines were were given a two-hour training program on communication methods, and were asked to encourage worker participation in the suggestion system through monthly production line "huddles", motivational worker meetings lasting three to five minutes.

I study a comprehensive set of worker and firm outcomes from these experiments by drawing upon a rich set of three information sources: intervention data (from the submitted, written suggestions), survey data (from baseline and endline surveys with workers and supervisors) and factory data (on production, quality and personnel). I estimate the short-run impact of the interventions on two outcomes categories: idea-driven voice, and a combination of worker well-being and firm performance. From perspectives of the workers and the firm, both sets of outcomes are important. First, giving workers a voice through encouragement of suggestions might enhance feelings among workers that the firm listens to them and cares about them. This empowerment might directly affect worker well-being and/or indirectly affect firm performance through higher motivation and effort of workers in the short-run – regardless of the quality of workers' suggestions. Second, collecting additional suggestions might be worthwhile in the short-run if their implementation has the potential to improve firm performance in the long-run, in line

[1]See Friebel et al. (2021); Hoffman and Tadelis (2020); Lazear et al. (2015)

with the Japanese production philosophy *Kaizen*.[2]

In examining issues on idea-driven voice, I find that the lunch intervention significantly increased the quantity and relevance of suggestions that are implementable in the firm. In fact, workers in the control group submitted very few suggestions – 48 suggestions from the control group compared to 1809 of the treatment group. However, the quality of the suggestions from the control group was 20% higher. (Quality was measured by an index of six criteria evaluated by an industry expert.) This result suggests a positive self-selection of high-quality submissions from workers in the control (non-lunch session) group. On the contrary, workers in the treatment group submitted suggestions of much larger breadth, covering diverse topics. By contrast, the huddle treatment had no effects on the quantity, quality or variety of suggestions. (Variety was assessed using a count of new categories over time, and a latent semantic analysis approach (LSA) that compared word vectors.) Moreover, in comparison to the lunch treatment, the huddles produced few suggestions on critical topics about supervisors' skills and behaviors.

In examining worker well-being and firm performance, I find that the lunch intervention improved worker well-being in the form of higher job satisfaction at the time of the endline survey, but it had no short-run effect on firm performance. The huddle treatment did not produce any effects on worker well-being and firm performance – either on its own or in combination with the lunch treatment. Overall, the lunch treatment was clearly beneficial, while the huddle treatment was not.

In analyzing the differences in quality and quantity of suggestions that emerged from the lunch control and treatment groups, I provide three arguments in favor of or trading off quality for quantity. That is, I argue that there is greater value from the lunch treatment group, whose suggestions varied in quality but were of greater quantity, than from the control group, whose suggestions were few but of higher quality. First, the lunch treatment increases the likelihood of receiving winning suggestions. Second, the larger breadth of topics addressed in submission gives the firm insight into what workers think and care about. Third, the lunch treatment makes workers better off regardless of the

[2]The Japanese concept of *Kaizen* (continuous improvement) covers different management techniques, in which everyone, workers and managers alike, are requested to consider continuous improvements as their permanent duty (Aoki, 1990).

quality of submitted suggestions. This could eventually lead to reduced turnover.

I examine the potential mechanisms by which the two managerial interventions have an impact. Based on survey outcomes, I find that the lunch intervention significantly decreases incentive-related constraints to idea generation ("Why should I bother?"). These effects take place despite winning suggestions not being implemented[3], suggesting that simply the act of exerting one's voice as part of the suggestion process makes workers more satisfied. By contrast, the huddle intervention imposes a cost, as it increases psychological safety-related constraints. It undermines workers' critical voice about their supervisors, as measured by a perceived deterioration in the voice climate. Workers exposed to treated supervisors feel less confident about sharing their opinions.[4] This finding is consistent with an effect defined in the literature as low 'psychological safety' (Edmondson, 1999; Edmondson and Lei, 2014).

I also analyze heterogeneous treatment effects and find that the lunch treatment successfully decreases incentive-related constraints for both workers with a below- and above-median experience index[5], but does not help to overcome psychological safety-related constraints for workers with a below-median experience index. I further provide evidence of a peer-mentor effect. Less-experienced workers benefit from being seated opposite a more-experienced friend during lunches.[6] The effect of friendship is identified because half of the workers in the lunch group were randomly assigned to be seated opposite a friend. This suggests that communication with high-experienced friends might help to overcome psychological safety-related constraints for workers with lower levels of experience.

This paper makes four primary contributions. First, it expands the growing literature

[3]Only 5% of surveyed workers notice the implementation of two non-winning suggestions: The instalment of a new fan and the repair of a broken floor tile. The firm itself made no major changes across the entire factory during the period of study, other than to adopt the measures that were part of the experiment (i.e., the lunches and huddles).

[4]Focus group discussions with supervisors delivered qualitative evidence that supervisors are worried about workers complaining about them to top managers, suggesting that supervisors could have indeed steered workers away from raising certain topics in line with Friebel and Raith's theory on skip-level communication (Friebel and Raith, 2004).

[5]The experience index is an Anderson (2008) index of the number of factories worked at, grade, tenure in the present factory, tenure in the industry and the educational level. I find similar results using the above-median number of processes that a worker can perform, measured by skill assessments.

[6]These results relate to those in Hahn et al. (2019) who find that the individual performance of lower-ability female children in Bangladesh improves once they study with peer friends. Similarly, Sandvik et al. (2020) find that being paired with a more-experienced co-worker during structured peer-to-peer lunches increases average sales revenue.

on the impact of management practices on firm-level outcomes (Bandiera et al., 2020; Bloom et al., 2018, 2013; Bruhn et al., 2018; Higuchi et al., 2017; Iacovone et al., 2018). The most closely related paper is the recent work by Sandvik et al. (2020) who find that interventions providing structured opportunities for workers to initiate conversations with peers foster knowledge flows among co-workers about each other sales techniques, and that monetary incentive systems based on joint output gains have no effects. While Sandvik et al. (2020) intervene with workers on the same hierarchical level, I focus on the effect of interactions between peers on bottom-up knowledge flows from production workers to higher-level managers. Plausibly, in my context there is much more to be learned from bottom-up knowledge flows than from peer learning due to the highly specialized and manual nature of each workers' tasks, such as certain sewing processes. In this setting, unlike that of Sandvik et al. (2020), idea generation is not achieved by efficiently connecting knowledge seekers and providers. Instead, it requires complementing different pieces of information, and conceptualizing combinations of knowledge; this is because the solution might be unclear to both knowledge seekers and providers.

Second, this paper relates to the literature on middle managers. For example, Lazear et al. (2015) and Hoffman and Tadelis (2020) find that middle managers are important for business performance, and Friebel et al. (2021) show that middle managers affect personnel turnover. By contrast, I empirically show that the involvement of middle managers does not improve the quality of vertical communication from workers to top managers. In fact, my findings show that their involvement leads workers to avoid critical topics about middle managers – particularly regarding concerns about their skills. I hereby touch upon the literature on worker voice in management science and organizational psychology that analyzes the conditions under which individuals feel empowered to speak up.[7] Many studies on worker voice[8] find that leader behavior plays an important role in motivating employees to speak up by reducing the fear of repercussions. I find that the involvement of middle managers degrades workers' perceived psychological safety; such psychological safety has been found to be a key antecedent of risk-taking behavior, such as suggesting

[7]See for example Coffman (2014) on idea contribution in gender incongruent areas, Rigtering et al. (2019) on managerial framing for suggestion systems, or Williams and Yecalo-Tecle (2019) on hostility by supervisors to new ideas from subordinates.

[8]See the literature overview by Morrison (2014) and Burris et al. (2018).

ideas (Edmondson, 1999; Edmondson and Lei, 2014).[9]

Third, this paper contributes to the literature on worker participation in firms' decision-making processes. Grant (1996) develops a knowledge-based theory of the firm, which assumes that knowledge resides within each individual. In his model, a participatory culture that allows knowledge transfer can increase both worker well-being and efficiency.[10] Several studies provide empirical evidence on the success of employee suggestion systems (Gibbs et al., 2017), and more generally, the positive impact of increased worker voice in firms' decision-making processes on economic performance and workers' welfare. Examples include worker representation in corporate boards (Harju et al., 2021), worker-manager safety committees (Boudreau, 2020), manager evaluations (Cai and Wang, 2020) or voice in the realm of both unsatisfactory work conditions (Cottini et al., 2011) and disappointing wage hikes (Adhvaryu et al., 2019). Existing studies in low-income countries have not looked at the effects of worker voice in the realm of improvements in both productivity and working conditions, and effective managerial methods in doing so.

Fourth, I contribute to the literature on the role of hierarchies in processing and communicating information. While hierarchies are beneficial for the coordination of activities (Bolton and Dewatripont, 1994), communication across hierarchies can be costly. Reasons for these costs have been extensively analyzed theoretically.[11] Empirical evidence on barriers to communication and how to overcome those is limited, however, particularly on communication across hierarchies. Previous literature analyzes, for example, incentive

[9]My paper also relates to the effect of "skip-level" communication, i.e., communication between subordinates and top management. Middle managers can have strategic concerns when communication between their subordinates and superiors is too open (Baird, 1977; Lillico, 1972). In fact, Friebel and Raith (2004) suggest in their theoretical model that skip-level communication may induce superiors to deliberately recruit unproductive subordinates, or abuse their personnel authority in other ways to protect themselves.

[10]His theory closely corresponds to Aoki's 1990 analysis of the Japanese corporation structure that decentralizes production decisions to workers. The Japanese concept of *Kaizen* (continuous improvement) covers management techniques such as quality circles, total quality management, kanban, or just-in-time measures. Another way to put their concept into practice is using suggestion systems to access and utilize the knowledge of individuals at low levels of the organization. For example, large industrial companies such as Toyota have successfully implemented suggestion systems. Sting et al. (2019) observe that more than 700,000 suggestions are implemented at Toyota per year.

[11]For example, costs may arise from various issues, including people's limited ability to transmit, receive or process information (Geanakoplos and Milgrom, 1991), the monitoring of subordinates (Qian, 1994) or the strategic use of information (McAfee and McMillan, 1995). More applicable to the context of this paper, communication costs can also arise from different organizational languages used by a principal and his agents, and a lack of a rich enough shared vocabulary (Arrow, 1974; Melumad et al., 1992; Prat et al., 2007; Weber and Camerer, 2003).

alignments between employees and owners (Atkin et al., 2017), adoption of innovative HR management practices (Gant et al., 2002), and co-location of workers (Battiston and Kirchmaier, 2018). I add to this literature by showing that peer-to-peer interactions enable production workers to conceptualize and sharpen their messages in an "organizational language" that top managers can understand and reward.

The paper proceeds as follows:. Section 2, explains the experimental design of the randomized controlled trial. Section 3 discusses the data collected, the outcome variables, and the empirical specification. Section 4 reports the main impacts on firm performance, worker well-being, and idea-driven voice. Section 5 presents the results on mechanisms. Section 6 concludes.

2 Experimental design

2.1 Setting and background

The study was conducted in a garment factory located in Gazipur, north of Dhaka in Bangladesh. The factory exports to large, international buyers, one of whom requested that the factory set up suggestion boxes on each of its sewing floors as a social compliance requirement. The factory agreed to collaborate on research to analyze how the suggestion submission process could best be facilitated. As a starting point, the factory placed suggestion boxes at easily accessible places for all workers on each floor. The factory made suggestion templates and pens available, and requested that workers submit suggestions about improvements for both productivity and working conditions. The suggestion template (shown in Figure C.4) contains pre-specified areas of improvement: absenteeism, defects, efficiency, machines, motions/waiting time, production flows/layout, safety, shipment, skills, wastage, work environment, or other. To incentivize extra effort into the process, financial rewards were set in place, offering to directly share with workers the payoffs that resulted from improvements they had suggested.[12] The factory introduced

[12]Even in the absence of implementation, financial rewards were paid out. In fact, none of the winning suggestions were implemented, and the factory itself made no major changes across the entire factory during the period of study, other than to adopt the interventions that were part of the experiment.

prizes for the best three ideas submitted per floor per month, with awards of 500 BDT (approximately 6% of the monthly salary of a helper and 5% of the monthly salary of an operator), 300 BDT, and 100 BDT. Floor managers initially introduced the suggestion box to all workers on the floor via the factory's public announcement system. Nevertheless, the suggestion box remained empty – even when monetary rewards of roughly double the original levels were put in place.[13] This suggests that knowledge flows across hierarchies are challenging, particularly in low-trust environments such as in Bangladeshi garment factories. The comments from a female Bangladeshi garment worker on the production floor, underscore the point. "Nobody gives importance to the floor girls, and they [supervisors/managers] won't even listen to us," she said.[14]

After tests of different interventions on one pilot floor failed to increase the number of submissions, the factory agreed to try a different approach. The factory agreed to roll out two managerial interventions: the structured co-worker lunch sessions, and the supervisor-led production line huddles (detailed in Section 2.2). The interventions involved 51 production lines on six production floors. The study population was composed of 1,588 workers who were registered on those production lines at the program start date.

2.2 The interventions

I conducted two cross-randomized managerial interventions, which were designed to shed light on different impediments to bottom-up innovation: effort and time costs, social costs, and the combination of both of these. Workers may face costs in terms of the effort and time needed to make suggestions. These costs may surface because workers may not be able to generate suggestions in an organizational language that managers can understand and reward (Weber and Camerer, 2003), or because workers face opportunity costs of time (Battiston and Kirchmaier, 2018), i.e. costs arising from investing time into offering suggestions instead of doing something else. Workers may also face social costs, arising from issues such as reputational concerns (Bursztyn and Jensen, 2017), fear of

[13]Offering rewards that were approximately 12% of the monthly salary of helpers and 10% of the monthly salary of operators) also did not lead workers to submit suggestions.

[14]The comment was made in a focus group discussion that took place on November 4, 2019, in one Bangladeshi garment factory with randomly selected workers.

signaling incompetence (Edmondson, 1999), and class differences between them and their managers, and altruistic feelings toward in-group co-workers (Hjort, 2014).

The first managerial intervention ("lunch treatment") targeted the costs of effort and time through structured peer-to-peer lunches offered. The lunches were offered to the treatment group of 1,588 workers, who were randomly assigned to be seated opposite friends or random workers, which we refer to throughout this paper as non-friend, to distinguish them from those they identified as friends.[15] The control group received a free take-away lunch (for fairness reasons), but no dedicated time for discussion, and no facilitated interaction. The second experiment ("huddle treatment") targeted social costs through a supervisor training program and line huddles with workers. Workers working on the 51 production lines assigned to the second managerial intervention were exposed to line huddles from trained supervisors – regardless of whether they were assigned to the first managerial intervention. The cross-randomized design allows me to categorize workers into one of the six groups depicted in Figure 1. For simplicity, I refer to workers assigned to the first managerial intervention as the "lunch group," and workers exposed to the second managerial intervention as the "huddle group." A more detailed description of each intervention is described in the following subsections.

2.2.1 Managerial intervention 1: Lunches

The idea behind the lunch managerial intervention is to test whether separating out time for workers to engage in discussions with co-workers about suggestions under the help of external facilitators improves the efficacy of idea generation. The intervention reduced opportunity costs of time (Battiston and Kirchmaier, 2018), arising from the fact that time spent thinking about and writing down suggestion cannot be devoted to other activities. I followed Sandvik et al. (2020) in providing free monthly lunches to a random subset of workers. The factory provided space and tables so that I could arrange for workers to interact with co-workers. .At the lunch, workers were able to ask questions to external facilitators, who provided suggestion papers and pens to each worker, and

[15]To clarify, a non-friend is a co-worker who has not been identified as a friend. A non-friend may be someone a worker has not previously met, or someone with whom a worker has had only limited interaction.

explained how to fill out the suggestion template. Suggestions could be handed over to external facilitators at the end of the lunch, or submitted any time into the suggestion box on the production floor. Workers were seated in the cafeteria on long tables with up to 18 seats. For logistical reasons, workers were partitioned into four subgroups based on their pre-assigned lunch hours. Every week, a different subgroup received lunch (in two cohorts from 12:00 to 12:45 p.m., and from 1:00 to 1:45 p.m.), such that each subgroup received one lunch per month over the course of three months (the length of a "round"). A detailed timeline can be found in Appendix Section C.2. Workers were informed about their assignment to a lunch session two to three days in advance because otherwise they often went home for lunch.[16]

For fairness reasons, the control group received a take-away lunch, which was handed out together with a suggestion paper and a short verbal instruction. However, the control group was not exposed to either the natural opportunity to interact with peers as part of a fixed seating arrangement, or to assistance from external facilitators. Workers in the control group were free to eat lunch wherever and whenever they wanted (e.g., inside the factory or at home). Most took the lunch home for dinner or for their families, and ate lunch as they usually did. They received a short instruction about how to fill out the suggestion paper, which was shorter in length than the one treated lunch workers received.

To measure the importance of trust in discussions, I randomized those receiving lunch to be seated opposite someone they had identified as a friend, or another randomly selected worker. I defined a friend as a production-floor co-worker with whom the worker most frequently discusses work-related topics; this information was taken from the baseline survey, which asked workers to name up to two co-workers with whom they communicate most frequently. By this definition, a friend is someone the worker can also trust in a discussion about suggestions. (Note: Throughout the paper, I refer to these co-workers with whom workers have frequent conversation as friends. However, when I asked workers to make their choices, I explicitly avoided using the term because the baseline survey revealed that many workers found it difficult to grasp the meaning of friendship per se.

[16]The baseline survey shows that more than 25% of the workers have lunches inside the factory less than every day, and that 15% of the workers never have lunch inside the factory.

In fact, the baseline survey showed that less than 1% of workers talk about private topics, and that more than 80% only talk about work-related issues with the person with whom they communicate most.) Some workers were able to coordinate the selections of closest non-friends in advance, and some were not able to do this. Thus, for simplicity, I impose symmetry on all friendships. That is, if A was mentioned as a friend of B, then B was also classified as a friend of A (even if A did not mention B as a friend). After two lunches, I re-randomized the seating of a random sample of one-third of workers, who were moved to be seated opposite someone other than their friends. This allows me to test whether being seated next to someone new after two lunches stimulates new ideas.

2.2.2 Managerial intervention 2: Huddles

In the second managerial intervention, supervisors working on randomly selected production lines were assigned to take part in a two-hour training program on worker communication methods at the start of the intervention; they were then instructed to conduct short motivational meetings ("line huddles") of two to five minutes with all workers on their line. They were told to explicitly encourage the workers to submit suggestions. The intervention was designed to test whether being encouraged by supervisors serves as an additional accelerator to make the first managerial intervention more efficient.

The supervisors received training on how to conduct the huddles. The training focused on worker communication methods, with both a motivational and a technical part. The motivational training gave supervisors information about the new suggestion system and their roles and responsibilities in the process. They received instruction about how to maintain good relationships with workers, how to encourage workers to submit suggestions (through on-the-job encouragement and line huddles), and how to give feedback. In the technical training, supervisors learned about how to generate solution-oriented suggestions; they were given insights about possible areas of improvement, and they were asked to share their experiences from the line huddles. The curriculum is shown in Table C.2.

To remind supervisors to conduct line huddles during the intervention period, I sent each treated supervisor three SMS reminders (see text in Figure C.3) with the request

to conduct a huddle within the next two to three days. After conducting a huddle, supervisors filled out a template (see Figure C.5), and submitted it to their industrial engineering (IE) manager for approval. The template was introduced as a commitment device, and to track whether huddles actually took place. The research team prompted IE managers to ensure that supervisors conducted the huddle, and did not just fill out the template with fake information.

2.2.3 Evaluation process

Over the course of the intervention, the research team collected suggestions during lunches or from the boxes. The research team digitized and anonymized the suggestions, which were then handed over to floor-level evaluation committees. The committees consisted of four to five nominated floor managers from different departments. To ensure an efficient evaluation process, each committee received a two-hour training with practical team exercises in January 2020 (see curriculum in Appendix Table C.2). Evaluation committees selected three winning suggestions per month, and submitted a nomination form to managers for approval (see template in Appendix Figure C.6). Approved suggestions were announced by floor managers. The winners were rewarded with prizes. The pictures and names of winners and the names of their supervisors were displayed on each floor's notice board. Pictures in Appendix Figure C.7 present the design of the display. The process was "semi-blind" in the sense that the workers' names were only anonymous before winning.

Simultaneously, an industry expert anonymously evaluated suggestions for research purposes.[17] Evaluations by the factory's evaluation committees and the industry expert show discrepancies (see Appendix Table C.4). For example, the evaluation committee gave awards to suggestions that the industry expert assessed as identifying a problem without offering a solution. These discrepancies might arise because evaluation committees may select not only implementable suggestions, but also suggestions containing ideas or observations that might not have immediately obvious solutions. Finally, suggestions and evaluations were translated into English.

[17] A detailed description of the quality index can be found in Section 3.2.1.

2.3 Randomized assignment

2.3.1 Randomization into lunch treatment

The randomization of 1,599 workers into lunch and non-lunch groups was implemented at the pair level (dyads of workers). I carried out the randomization in a six-step approach that I describe briefly here and in detail in Appendix Section D. I split the sample into two groups: workers who named (or were named by) at least one friend[18] (Sample 1), and those who did not name (or were not named by) a friend (Sample 2). I then pair workers in Sample 1 with a friend, and those in Sample 2 with a random worker, a "non-friend." I then randomize the dyads of friends and non-friends into treatment and control groups.[19] Further, I re-balance treatment assignments to achieve balance on the maximum experience index per dyad across treatment groups because I expect this to drive the quality of suggestions.[20] I also re-balance on the percentage of treatment dyads across lines to avoid the potential for the effects of the second managerial intervention to be driven by the number of workers in the lunch group participate in one huddle. Then, to allow for within-sample comparisons (within the group of workers with friends), I randomly select 50% of the friend dyads within Sample 1 to be re-paired with a non-friend on the same floor. Finally, after lunch Round 1 and Round 2, I randomly select one-third of all pairs assigned to lunch to be seated opposite a new non-friend, as shown in Table D.8. The final number of workers per treatment assignment is shown in Appendix Table D.6. The seating order of dyads along the tables in the cafeteria is randomly chosen.

2.3.2 Randomization into huddle treatment

The randomization into the huddle treatment took place at the line level, with 51 production lines randomly selected into huddle and non-huddle lines, using the production floor as a stratification variable. The randomization was carried out 1,000 times. For

[18]Friend is defined as one of the two co-workers a worker talks to most, according to the baseline survey.

[19]I stratified by sample, production line, and an indicator equal to one if one worker per dyad had already been asked extended survey questions. While dyads in the treatment group were seated opposite each other, the dyadic structure did not play any role in the control group.

[20]The experience index reflects a worker's maturity and skillset, and is measured by an Anderson (2008) index of the number of factories worked at, grade, tenure in the factory, tenure in the industry, and educational level.

each iteration, I identified the minimum p-value of two-sided t-tests of mean equality between treatment and control lines per balancing variable. The variables are the number of workers per supervisor, efficiency of the production line as measured by the output minutes divided by the number of workers per line, alteration rate as measured by the number of alterations divided by the number of pieces checked per line, and absenteeism rate as measured by the number of workers absent divided by the number of total workers per line (present and absent). T-tests are conducted separately for production lines on three low-performing and three high-performing floors, where performance is based on a ranking of all floors according to efficiency, alteration rates, and absenteeism. This step was undertaken to avoid a treatment assignment of only low-performing lines on some floors and only high-performing lines on other floors. Out of the assignments with a minimum p-value of > 0.20, I chose one at random.

2.3.3 Selection of survey participants

The research team attempted to survey all workers included in the experiment at baseline and endline with a short set of questions. A subset of workers was selected for an extended version. The extended baseline survey was conducted in part at the factory and in part on the phone. I randomly selected three operators and one helper per production line, or 204 workers on the 51 production lines. Because of COVID-19, the extended version of the endline survey was conducted entirely on the phone. I selected at random from each production line[21] two workers from the control group, one worker from the lunch group seated opposite a friend, and one worker from the lunch group seated opposite a non-friend. Because it was much harder to reach workers at endline, I randomly selected one additional reserve worker per initially selected worker on the same production line and with the same treatment assignment.[22] I surveyed all supervisors on the six participating production floors.

[21] A worker's production line was defined as the line where she worked most of the time between January and March 2020. The definition was based on data on line movements provided by the factory. If there were not enough workers left per group per line due to high turnover, I added reserve workers who were on the respective line at baseline. This was the case for two lines for the lunch group of workers seated opposite friends, and for three lines for the lunch group of workers seated opposite non-friends.

[22] Enumerators were asked to call reserve workers only if the initially selected worker could not be reached after nine attempts at different points in time. In all, 68% were surveyed from the initial list in the control group, 74% in the treatment group, and 76% in the non-friend treatment group.

3 Data collection and measurement

3.1 Data sources

The analysis draws upon four main sources of information: Intervention output data, administrative data, survey data, and focus group-generated data. First, I obtained the intervention outputs, the submitted suggestion papers filled out by workers, huddle forms filled out by supervisors, and nomination forms filled out by the evaluation committees. Second, I used various factory datasets maintained by different departments. The production department provided daily production data including efficiency and quality information. The HR department provided monthly salary data (information about the workforce), daily job card data (information about times workers entered and left the factory) and skills data (information about each process a worker can perform). The IE department provided line movement data with the daily working location of each worker. Due to the incompleteness of the data,[23] I drew upon intervention and survey data to complement the line location of workers (see description in Appendix Section B). Third, the analysis uses survey data of line operators and supervisors on treatment and control lines collected at two points in time: shortly before the experiment in early December 2020 (in-person and phone baseline survey), and shortly after the experiment in mid-March 2020 (phone[24] endline survey). Both the baseline and the endline surveys of line operators consist of a short version for all workers, and a long version for randomly selected workers. For supervisors, there was only one version. Finally, I organized three focus groups with workers, lower/middle managers, and higher-level managers (separately) in November 2019 at the end of the pilot experiment. The focus groups were conducted to provide a deeper understanding into the mechanisms through which the treatment was expected to have an impact. Further detail on the four data sources can be found in Appendix Section E.

[23] 60% of the days during the intervention period are missing, and 159 out of 1,599 workers are not included in the data at all.

[24] The endline survey was conducted on the phone only, as enumerators were not able to visit the factory due to Covid-19.

3.2 Primary and secondary outcome variables

3.2.1 Primary outcomes

I analyze two groups of primary outcome variables, which I expect to change in the short run as a result of the intervention. I estimate treatment effects on idea-driven voice (i.e., the submitted suggestions of workers), and on worker well-being and firm performance. These two groups of outcomes are important because both, either together or independently, could be beneficial for the firm and workers both in the short and long runs. For example, giving workers a voice through encouragement of suggestions might make them feel more listened to and cared for – even if their suggestions were irrelevant, not awarded with a prize and not implemented. This empowerment might directly affect worker well-being and/or indirectly affect firm performance through higher motivation and effort of workers in the short run. At the same time, even if worker well-being and firm performance were not affected, collecting additional suggestions might be worthwhile in the short run if implementation of workers' ideas has the potential to improve firm performance in the long run.[25] This concept is in line with the Japanese production philosophy *Kaizen*.

i. Idea-driven voice. I analyze treatment effects on three outcomes: quantity, quality and variety of suggestions. The quantity of suggestions is the total number of suggestions per worker per round, where round is the time in which a treated worker received one lunch. The quality of suggestions is measured with an index score on a scale of one to ten. I use two versions of the quality index to take into account the trade-off between quantity and quality. The first considers only the quality of *submitted* suggestions, with the index starting at one point and workers not submitting any suggestion coded as missing. The second codes workers not submitting any suggestion as zero, so is defined for all workers. In the second version, a bad suggestion is considered a better outcome than no suggestion at all. Points are based on an expert's anonymous evaluation, in which the identity of the author of the suggestion is not displayed. Experts first evaluate the type of the suggestion out of six possible types: They categorize the suggestion

[25]Long-run effects are not part of the analysis in this paper.

according to whether it is incoherent, offers general observation but does not identify a problem, identifies a problem (communicating this either poorly or clearly), or offers an idea (again, communicating this either poorly or clearly). Ideas are then further assessed on the impact that they would have if they were implemented according to the six components that indicate the extent to which a suggestion is likely to benefit workers and/or the firm. I examine the potential impact of the idea on two worker components: the level of worker impacted, and the effect on worker well-being. I also examine four firm components: the likely impact on efficiency, the difficulty of implementation, the financial resources required to implement the idea, and the nature of productivity effects (direct or indirect through improved motivation of workers). More details on the quality evaluation can be found in Appendix Section G.

The variety of suggestions is measured as the number of problems and ideas per round (suggestions of type three to six) that have not been mentioned by the respective worker in any previous round. Variety can only be measured for workers who submitted more than one problem and/or idea. To identify whether suggestions have been mentioned previously, suggestions are classified according to three different levels of granularity (from least to most granular) according to category, subcategory, and suggestion. The category used to classify suggestions does not necessarily match the area of improvement that a worker might have indicated on the suggestion form. Appendix Figure G.9 provides a detailed overview of the classifications. I analyze the variety at the most-granular (suggestion) level. As one robustness check, I use the less-granular category and subcategory measures (see detail in the appendix). As a further robustness check, I generate an additional measure of variety based on latent semantic analysis (LSA), a technique in natural language processing that allows comparisons of word vectors (Schwarz, 2019).

As an additional outcome, I measure the content of suggestions. I generate an indicator equal to one if the suggestion is classified as a certain category that corresponds to one of the following four content areas: supervisor skills, vacation/attendance bonus, work environment, or production (see Appendix Figure G.9).

ii. Firm performance and worker well-being. I analyze treatment effects on four families of outcomes, related to workforce, production, job satisfaction, and mental health. Workforce and production outcomes are directly related to firm performance. Job satisfaction and mental health outcomes are directly related to worker well-being. To reduce the number of hypotheses tested, I use a standard "omnibus" approach; I group outcomes into families of outcomes, construct an index for each family, and test whether each index is affected by the treatments. Thus, I first re-code all outcomes such that higher values correspond to a similar direction of change within an index. I then standardize the outcomes to have a mean of zero and standard deviation of one. I then construct a weighted index of outcomes using the method of Anderson (2008). I calculate the average of the standardized constituent outcomes weighted by the inverse covariance matrix. The family indices are comprised of the following constructed variables (listed in more detail in Table G.9):

- **Workforce index:** Absenteeism, punctuality, and working hours
- **Production index:** Efficiency, alterations, and reject rates
- **Job satisfaction index:** Job satisfaction and support, turnover intentions, and work attitudes
- **Mental health index:** Feelings at work, general anxiety disorder and workplace harassment

3.2.2 Additional secondary outcomes

I also test a wide range of secondary outcomes to explain potential mechanisms behind the results. Secondary outcomes are based on baseline and endline survey questions. Appendix Figure G.10 displays index components for secondary outcome variables. Indices are also constructed using the method of Anderson (2008).

3.3 Empirical specification

To estimate treatment effects on the primary and secondary outcome variables, I calculate the intent-to-treat (ITT) estimates using four simple regression models. I also calculate

local average treatment effects (LATE), as described in the Appendix Section F.

The first ANCOVA (Analysis of Covariance) specification serves to test the effect of the first and second managerial interventions separately:

$$Y_{i,l,t} = \alpha + \beta_1 T_{i,l} + \theta Y_{i,l,t=0} + \gamma_{i,l} + \epsilon_{i,l,t} \tag{1}$$

where $Y_{i,l,t}$ is the outcome variable of interest for worker i on line l at time t. $T_{i,l}$ is an indicator for random assignment to the treatment, which is either the first managerial intervention $Lunch_{i,l}$ or the second managerial intervention $Huddle_{i,l}$. $Y_{i,l,t=0}$ is a control for the baseline value of the outcome variable. $\gamma_{i,l}$ is a stratum indicator (experience index), and $\epsilon_{i,l,t}$ is the residual. β is the coefficient of interest.

The second regression model serves to test interaction effects:[26]

$$Y_{i,l,t} = \alpha + \beta_1 Lunch_{i,l} + \beta_2 Huddle_{i,l} + \beta_3 Lunch_{i,l} * Huddle_{i,l} + \theta Y_{i,l,t=0} + \gamma_{i,l} + \epsilon_{i,l,t} \tag{2}$$

where $Lunch_{i,l}$ is an indicator for random assignment to the lunch treatment, $Huddle_{i,l}$ is an indicator for random assignment to the huddle treatment, and $Lunch_{i,l} * Huddle_{i,l}$ is an indicator for random assignment to both the lunch and the huddle treatment. The notation for equation 2 is otherwise analogous to that for equation 1. In this specification, β_1 is the estimated treatment effect of the lunch treatment, β_2 is the estimated treatment effect of the huddle treatment, and β_3 is the additional effect of receiving both treatments.

The third regression model allows me to test for heterogeneous treatment effects:

$$Y_{i,l,t} = \alpha + \beta_1 T_{i,l} + \beta_2 HET_{i,l} + \beta_3 T_{i,l} * HET_{i,l} + \theta Y_{i,l,t=0} + \gamma_{i,l} + \epsilon_{i,l,t} \tag{3}$$

where $HET_{i,l}$ is an indicator for the above-median baseline value of an interaction variable. The notation for equation 3 is otherwise analogous to that for equation 2. In this specification, β_1 is the estimated treatment effect on workers with a below-median

[26]Because the lunch treatment started on December 26, 2020, and the huddle treatment started later, on January 14, 2020, I restrict the data to all observations after the huddle start date in regression models that test the effects of the huddle treatment, or the effects of the lunch and huddle treatments jointly. In regression models that test the treatment effects of the lunch treatment only, I include all observations.

baseline value of the interaction variable, $\beta_1+\beta_3$ on workers with an above-median baseline value, and β_3 is the estimated difference between these two treatment effects. The fourth regression specification serves to test effects on lunch treatment days:

$$\begin{aligned} Y_{i,l,t} = \beta_0 + \beta_1 td_0 + \beta_2 td_1 + \beta_3 td_2 + \beta_4 td_3 + \beta_5 td_4 + \beta_6 td_5 + \beta_7 td_6 + \beta_8 td_{-1} \\ + \beta_9 td_{-2} + \beta_{10} td_{-3} + \beta_{11} td_{-4} + \beta_{12} td_{-5} + \beta_{13} td_{-6} + \theta Y_{i,l,t=0} + \delta_t + \epsilon_{i,l,t} \end{aligned} \tag{4}$$

where td_0 is an indicator for the lunch treatment day. A positive number in the subscript indicates leads (number of days after the treatment day), and a negative number in the subscript indicates lags (number of days before the treatment day). β_2 to β_7 are the treatment effects on the first day up to the sixth day after the treatment day, whereas β_8 to β_{13} are the treatment effects one to six days before the treatment day.

Standard errors are clustered at the pair level in regression models that test effects of the lunch treatment only. Standard errors are clustered at the line level in regression models that test effects of the huddle treatment, or interaction effects. All regressions include floor fixed effects to control for any time-invariant differences across floors. Time fixed effects are also included unless the data are collapsed at the individual level.[27]

I report a p-value for each family of outcome variables. I also separately report effects on individual variables. To account for the fact that I am testing multiple hypotheses, I consider more conservative p-values. For each individual variable (i.e., a sub-component of primary outcomes), I report a p-value and false discovery rate (FDR) q-values calculated within the family to control for the expected proportion of rejections that are false positives. I do not correct for multiple hypotheses testing when looking at secondary outcomes because this analysis is informative for exploratory analysis of additional impacts, robustness checks, and mechanism analysis – not the main impact.

[27]Time fixed effects are in rounds for innovation-related outcomes, in weeks for production-related outcomes, and in days for workforce-related outcomes. Regressions on index variables based on survey data (i.e., job satisfaction- and mental health-related outcomes), and all secondary outcomes do not include time fixed effects.

3.4 Integrity of the experiment

3.4.1 Baseline Balance and Summary Statistics

I confirm the validity of the randomization by performing a balance test, results of which are shown in Appendix Table I.11 for primary outcomes, and in Appendix Table I.12 for secondary outcomes and characteristics of workers. Only two of the twelve primary outcome variables (absenteeism and general anxiety) in Appendix Table I.11, and none of the secondary outcome variables or characteristics of workers in Appendix Table I.12 are different at the 10% level. A joint orthogonality test of worker characteristics for each treatment and outcome group shows that overall balance cannot be rejected. A few characteristics are worth highlighting (see Appendix Table I.12). Around 80% of the sample workers are female, and the average age is 26; by contrast 28% of the supervisors are female, and the average age is 30. Workers have completed on average of eight years of schooling, which implies that they should be able to read and write. In Bangladesh, students can go to a senior school and then to a high school after nine years of schooling. Less than 2.5% of workers in the sample have completed only two years of schooling and, hence, might not be able to write suggestions themselves. For these workers, external facilitators or supervisors could be asked to write down suggestions. Supervisors have also completed on average eight years of schooling, which is not surprising given that most supervisors are recruited from a pool of experienced operators. Workers are able to perform on average five different sewing processes, and work on average on 3.5 different lines over a three-month period.

3.4.2 Compliance

Non-compliance is only one-sided for the lunch treatment and not very strong. Control workers never receive any lunch, and less than 10% of workers assigned to the lunch treatment did not receive any lunch. Therefore, the assigned lunch and actual lunch variables are almost identical, and LATE and ITT effects are very similar.

By contrast, non-compliance for the huddle treatment is two sided, and stronger for two reasons. supervisors incorrectly conducting or failing to conduct line huddles; and

workers extensively moving across lines. Regarding supervisor conduct, an evaluation of the line huddle forms indicates that non-compliance is low (see Appendix Table C.3). The evaluation shows that 92% of the treatment lines (23 out of 25 lines[28]) correctly conducted at least one huddle in January 2020; it shows that 15% of the control lines (4 out of 26 lines) incorrectly conducted a huddle by mid-February 2020, indicating some minor non-compliance on control lines. The larger source of non-compliance comes from the fact that workers moved across lines. Workers were assigned to an average of 3.4 lines during the huddle treatment window, and the average maximum number of days per worker per line during the 90-day period between January and March 2020 is 56 days.[29] I cannot rule out that workers' line movements are endogeneously related to unobservables. For example, workers move more often across lines if they can perform multiple processes.[30] These workers might also submit more and better suggestions, implying that workers' line movements are not exogeneous to treatment. Thus, I rely on ITT effects when interpreting results. LATE (IV) effects on primary outcomes are shown in the appendix.

3.4.3 Attrition

Due to COVID-19, the endline survey had to be conducted by phone. As a result of the pandemic, the factory closed, and most workers went home to their families, making it much more difficult to reach them. The survey team made a great effort to follow up, attempting to contact each worker up to nine times. In total, 73% of the sample was re-surveyed at endline. Of concern is whether treatment is correlated with attrition. I test for this in Appendix Table I.14 by regressing a dummy variable indicating if the worker was not found at endline on treatment indicators. I find no significant differences in attrition rates across treatment arms. For supervisors, I received an updated list with phone numbers of all supervisors. I ensured that at least one supervisor was surveyed per line, independent of whether he had been surveyed at baseline.

[28]Two of 25 lines did not conduct a huddle by January 2020: One treatment line conducted its first huddle mid-February, and one failed to conduct any huddle in the entire treatment period.

[29]See Appendix Figure H.12 for the distribution of the number of lines per worker during the intervention period, and Appendix Figure H.13 for the distribution of the number of days working on the same line.

[30]A simple OLS regression in Appendix Table I.13 shows that line movements are increasing (column 1) and the maximum number of days per line are decreasing (column 2) in the number of sewing processes a worker can perform with p-values of 0.11 and 0.085, respectively.

4 Main outcomes

4.1 Impact on idea-driven voice

I begin by examining the impact of idea-driven voice on the quantity, quality, and variety of suggestions. The results are shown in Table 1. I first describe the effects of the lunch intervention (the structured peer-to-peer lunches), and I then analyze additional effects of the huddle intervention (the supervisor training and line huddles). The main results hold using LATE (IV) specifications as shown in the Appendix Table J.17.[31]

I find that the lunch intervention significantly increases the quantity and relevance of suggestions that are implementable in the firm. Column 1 shows that workers who were not assigned to lunch have a 3% probability of submitting suggestions, and this leaps to 59% for those in the lunch group. Figure 2 illustrates this result by showing the large difference in the number of submitted suggestions per round. The total number of suggestions from the lunch group over the course of the intervention period is 1,809, compared to 48 from the non-lunch group.

While the quantity of submissions from the lunch group eclipses those from those who did not participate in the sessions, the intervention has the opposite effect on the quality of suggestions. The quality is on average 1.2 points lower in the lunch group compared to the non-lunch group (as indicated by the significantly negative coefficient in column 7). This represents a 20% lower quality level between the lunch and non-lunch groups (as measured by comparing the mean quality levels). Figure 3 illustrates this result by showing the distribution of the quality index of suggestions with a larger tail at the top end for the non-lunch group, stemming from a higher quality related to firm components as shown in Table J.21. Overall, this result suggests that there is a positive self-selection of high-quality submissions from workers not assigned to the lunch treatment. It is more costly for them to submit suggestions in terms of required time and effort, so the cost acts as a filter.[32]

[31] Results are also robust to excluding randomization of Sample 2 (see Appendix Table J.15 for regression results, and Section D for explanations of the sample composition).

[32] Similar results hold using a quality threshold instead of the quality index. As a quality threshold, I use an indicator equal to one if the suggestion is an idea (type five or six). Appendix table J.20 shows the results.

Section G described a second version of the quality index, in which suggestions of low quality are considered as better than no suggestion at all (i.e., non-submissions are coded as zero instead of as missing). By this measure, the quality index of the lunch group is 24% higher than of the non-lunch group. Figure H.17 illustrates the difference between the two measures of quality, with higher quality of the non-lunch group in the upper right graph, and higher quality of the lunch group in the upper left graph.

I also find that workers in the lunch group submitted suggestions of much greater breadth, covering a wider diversity of topics than the suggestions from the non-lunch group. In fact, the non-lunch group mainly submitted suggestions about leave (vacation) practices. In the non-lunch group, some categories – such as work environment/facilities, salary, motion/waiting time, efficiency, and health and safety – were not addressed at all (see significantly negative coefficients in Panel (a) of Appendix Table J.23). Because workers in the non-lunch group never submitted more than one suggestion, the variety of suggestions from individuals cannot be compared between the lunch and non-lunch group.

The effects of the huddle intervention are less promising. Neither the quantity nor the variety of suggestions changes (see columns 2, 5, 8 and 10 of Table 1). Results on variety are robust to different specifications of variety based on categories and subcategories, or a latent semantic analysis approach as shown in the Appendix Table J.22. There are also no interaction effects between line huddles and lunches (see columns 3, 6 and 9 of Table 1). The negative effect on quantity in column 3 is even marginally significant. However, I find more robust evidence that the huddle treatment influences the content of suggestions. The huddle and non-huddle groups submit a similar number of suggestions, but the huddle group submits significantly fewer suggestions about supervisor skills (see Figure 4, and also see column 1 of Appendix Table J.23 panel B). The category of supervisor skills involves suggestions about two subcategories: behavior/tasks (e.g., no scolding, managing work allocations), and supervisor gender (e.g., avoiding female supervisors). Overall, these results suggest that the encouragement of supervisors does not serve as an additional accelerator to make peer-to-peer lunches more efficient, and that the presence of supervisors even undermines workers' critical voice about supervisors' skills.

4.2 Impact on worker well-being and firm performance

The impact on firm performance (represented by the workforce index and the production index) and on worker well-being (represented by the job satisfaction index and the mental health index), is shown in Table 2. The main results hold, using LATE (IV) specifications as shown in the Appendix Table J.18.[33]

I find no short-run effect of the lunch intervention on firm performance, but I find significantly positive effects on worker well-being, particularly on the job satisfaction index with an estimate of 0.2 in column 5 and of 0.3 in column 7 once both treatments are included. A closer look at the components of the job satisfaction index in Table J.19 shows that the positive effect is driven by significantly improved job attitudes in column 10. This result survives multiple testing correction with a q-value of 0.055. I calculate q-values, the false discovery rate (FDR)-adjusted p-values, following the method of Benjamini et al. (2006). The job attitudes sub-index includes questions about the extent to which workers do the best job they can, and whether workers understand the importance of punctuality. A respect for punctuality is an important aspect of factory work, given the sequential nature of the production process. Workers might not be able to start working if another worker on the line, who has to perform a preceding sewing step, does not show up on time. Coefficients on the job satisfaction sub-index (column 8) and turnover intentions sub-index (column 9) are positive, but not significant. The workplace harassment sub-index, a component of the mental health index, also has a significantly positive effect of 0.2. This index includes questions about the frequency of the use of abusive language or how often uncomfortable situations arise at work. One must bear in mind that the endline survey was conducted in the midst of the outbreak of COVID-19; this explains why the control mean of the job satisfaction and mental health index is negative at endline (see Figure H.15). The lunch intervention resulted in a much less severe deterioration of the job satisfaction index in the treatment than in the control group. This makes the result of improved worker well-being for the lunch treatment even more promising; positive effects on other sub-indices, such as general anxiety, plausibly might have been stronger in the

[33]Results are also robust to excluding the randomization of Sample 2 (see Appendix Table J.16 for regression results, and Section D for explanations of the sample composition).

absence of the pandemic effects. For example, the variance of general anxiety in the control group increases by 79% between baseline and endline, suggesting that statistical power might be lower in this measure.

I also examine effects of the lunch treatment at the daily level by including leads and lags of the treatment day in the regression specification. These effects can only be estimated for variables available at the daily level (i.e., the workforce index components). I find that workers were significantly less likely to have been absent on days when they were assigned to receive a free lunch and on the following day (see plot of coefficients in Appendix Figure H.14). This result illustrates that workers make sure not to be absent so they can receive a free lunch from the firm, and so they can participate at the suggestion process, which could make them feel valued and cared for. There are no effects on punctuality or working time on lunch treatment days.

The huddle treatment has no significant effects on any of the four primary outcome variables. Looking at the sub-components in Table J.19, only two of 13 measures are significant, both at the 0.10 level. Working time increases and alterations decrease, both marginally significantly. Moreover, the coefficients on both working time and alterations are negligibly small; the results do not survive multiple testing correcting.

4.3 Evaluation of quantity-quality trade-off

In evaluating the lunch intervention, the firm is confronted with a quantity-quality trade-off to either get a higher quantity of suggestions with a larger variance in quality from the lunch group or a lower quantity of suggestions with a higher quality from the non-lunch group.[34] I provide three explanations to conclude that the latter is more beneficial for the firm and the modest costs to evaluate suggestions are offset.[35] First, the lunch treatment

[34]This trade-off is captured by using two versions of the quality index as described in Section 3.2.1. The first version codes non-submissions as missing, suggesting a preference of quality over quantity. The second version codes non-submissions as zero, suggesting a preference of quantity over quality while simultaneously considering quality, i.e. suggestions of low quality are considered as better than no suggestion at all, but suggestions of high quality are also considered as better than low-quality suggestions.

[35]It takes approximately 30 seconds to assess the type of a suggestions and additional 60 seconds to further evaluate ideas. Assuming that each worker per floor submits one suggestion, i.e. 250 suggestions, and 30% of those are ideas, the floor-wise evaluation committee needs to invest only about 3.3 hours per month.

increases the likelihood of receiving winning suggestions: 88% of the suggestions judged to be best (blind to treatment status) were from the lunch group. The factory awarded five suggestions from the non-lunch group and 38 suggestions from the lunch group with prizes. Winning might make workers feeling better, and awarded ideas might be more efficiency enhancing once implemented (not assessed in this paper). Second, the lunch treatment gives the firm insights into what workers think and care about: More than 60% of the lunch group has submitted at least one suggestion, and a larger breadth of topics is covered (as outlined in Subsection 4.1). Implementing suggestions on those topics could eventually lead to productivity improvements in the long-run. In fact, some workers already noticed the implementation of a few suggestions at endline, such as the instalment of new fans or the repair of broken floor tiles - even though these suggestions were not announced as winning suggestions. Third, the lunch treatment improves worker well being. It makes workers significantly better off in terms of job satisfaction regardless of the quality of submitted suggestions (as outlined in Subsection 4.2). This could eventually lead to reduced turnover. The highest turnover is usually around the Muslim holiday Eid al-Adha[36], which could not be measured as part of this project due to Covid-19. The lower absenteeism on assigned lunch treatment days is suggestive of workers enjoying to be involved in the suggestion process and to receive a free lunch gift from the firm. Hence, I conclude that suggestions of low quality are better for both the firm and workers than no suggestion at all.

5 Mechanisms

In this section, I look at potential mechanisms that can explain results on primary outcome variables. I first analyze how workers' perceived constraints to raise their voice have changed, using survey outcomes. I then look at heterogeneous treatment effects, deriving results at the extensive and intensive margin. Here I focus exclusively on the lunch group, as the non-lunch group barely submitted any suggestions at all.

[36]There are two key Eid's (Celebration Festivals) in Islam: Eid-ul-Fitr, which signifies the completion of the Holy Month of Ramadan; and Eid-ul-Adha, the greater Eid, which follows the completion of the annual Hajj pilgrimage, at the time of Qurbani (sacrifice) (BBC, 2019). They are celebrated around May and August, respectively, although the larger turnover is usually after the second Eid in August.

5.1 Changes in perceived constraints

The communication of ideas from production workers to managers in hierarchical organizations can be limited as the result of costs of effort and time, and social costs. As outlined in Section 2.2, workers may face costs of effort and time to develop ideas, arising from, for example, the inability to generate suggestions in an organizational language that managers can understand and reward (Weber and Camerer, 2003), or opportunity costs of time (Battiston and Kirchmaier, 2018). Workers may also face social costs that limit the diffusion of information, arising from, for example, reputational concerns (Bursztyn and Jensen, 2017), fear of signaling incompetence (Edmondson, 1999), or class differences between managers and workers, and loyalty toward in-group co-workers (Hjort, 2014). I test whether treated workers and supervisors change their perceptions about these costs. In particular, I analyze responses to survey questions related to the following constraints:

- Incentive-related constraints, which increase costs of effort and time and/or reduce perceived benefits (targeted by lunch treatment)
  - Feeling of futility: "I don't think it's worth submitting suggestions" (futility index)
  - Lack of time: "I don't have the time to make suggestions" (time index)
- Psychological safety-related constraints, which increase social costs (targeted by huddle treatment)
  - Lack of knowledge: "I don't know what to suggest" (knowledge index)
  - Insecurity: "I don't dare to make suggestions" (security index)

The first managerial intervention (lunch treatment) targets costs of effort and time by reducing incentive-related constraints, and the second managerial intervention (huddle treatment) targets social costs by reducing psychological-safety related constraints. I present ITT effects of both interventions on all constraints in Table 3 for completeness.

5.1.1 Perceived decrease in incentive-related constraints

I begin by explaining the effects of the lunch intervention (see column 2 in Table 3). Why do workers assigned to lunch submit more suggestions? I examine the responses that relate to workers' perceived incentive-related constraints, targeted by the lunch intervention. While I do not find any changes in the time index, I do find significant changes in the futility index. The time index captures the frequency of idea generation. Exemplary survey questions are "How often do you think about ideas over lunch?" or "To which extent is time a constraint for idea generation?". While opportunity costs of time are reduced by design of the experiment (i.e., time is set aside for workers to jointly discuss suggestions), workers themselves do not recognize time as a constraint for idea generation. The futility index captures the workers' perceived probabilities of winning (i.e., their beliefs), which could drive results in line with the theoretical framework. The survey asked workers to predict the number of suggestions that they think they will submit the next month, and to estimate the likelihood that they will win, and the likelihood that their own idea will be implemented (detailed overview of all variables in Appendix Figure G.10). The coefficients from regressing the lunch-treatment indicator on those survey variables are significant, as shown in Appendix Table K.25. For example, the coefficient on the belief about winning and the belief about the implementation of a worker's own idea are significant at the 1% level and large in size (0.38 and 0.40 standard deviations, respectively).[37] The results show that workers have developed significantly greater aspirations to submit suggestions and to win a prize, and are much more confident of their own contribution to the suggestion system. This is also what external facilitators experienced during lunch sessions. Initially, many workers found it challenging to come up with suggestions, and they were concerned about not being able to offer anything. Only after external facilitators constantly and clearly explained the task and intensively motivated workers did the workers develop an awareness that they actually had something

[37]Furthermore, I find significantly positive effects on individual index components of the voice climate index related to the ease of submission, e.g., feeling comfortable submitting ideas, or having the ability to openly state opinions (see Appendix Table K.26). However, even though lunch workers receive instructions from external facilitators, and discuss ideas with co-workers, the perceived ease of generating ideas, as measured by the knowledge index, does not improve. Workers do not find it easier to think about problems and solutions, suggesting that the lunch treatment does not seem to have improved workers' perceived intellectual abilities to generate suggestions.

to contribute. External facilitators were explicitly instructed not to provide workers with content for suggestions. This instruction was provided to rule out the possibility of technical help driving any changes in aspirations.

These effects take place despite workers experiencing no actual changes in firm's policies or of managers' behaviors. The increase in job satisfaction as outlined in Section 4.2 seems to be driven by the act of exerting one's voice, regardless of the implementation of suggestions.[38] Several findings support this argument in line with what has been causally explained in the literature (Adhvaryu et al., 2019).[39] When asked whether managers/supervisors had implemented any ideas from workers over the previous six months, 95% of all surveyed workers responded with "nothing/don't know." (see Appendix Figure H.19). Only 5% noticed the implementation of two suggestions at endline: The instalment of a new fan and the repair of a broken floor tile. Both suggestions were not awarded with prizes. Moreover, neither the communication practices index (see Appendix Table K.27) nor the security index measuring the voice climate and leadership efficacy (see Table 3) show significant changes. Workers did not experience any actual changes in workplace practices or culture, and continue to believe that nothing will change. This is not surprising given that none of the winning suggestions were implemented, and the firm itself made no major changes across the entire factory during the period of study, other than to adopt the measures that were part of the experiment (i.e., the lunches and huddles).

Overall, I conclude that the lunch treatment has successfully illustrated workers that the firm really seeks their ideas ("Why should I bother?"). It has addressed impediments to idea generation by changing workers' perceived benefits and aspirations to participate in the suggestion process.[40] These effects take place despite the lack of actual implemen-

[38]In my theoretical framework (described in Appendix Section A), this result corresponds to changes in the intrinsic motivation of workers to generate ideas, $M(e_i; \lambda_i, N_{ij})$, which captures the fact that workers derive utility from thinking about ideas, regardless of winning or implementation of ideas. For example, they derive utility through the enjoyment they experience by being listened to and by being cared for, and/or they gain inner satisfaction from thinking about ideas individually or jointly.

[39]Adhvaryu et al. (2019) show that giving randomly selected workers the opportunity to participate in a feedback survey after a disappointing wage hike reduced turnover. Participants were told that the firm would not receive the feedback provided; thus, the authors conclude that the act of being asked for one's opinion drives the result, rather than any changes in firm policy or managers' behaviors.

[40]I examine a number of different alternative hypotheses to explain the impact of the lunch treatment, but I do not find any compelling evidence for any of these potential explanations. By design, I rule out gift-exchange as a core mechanism (Dellavigna et al., 2016, 2018; Gneezy and List, 2006); this is because the control group also received an individual take-away lunch with a suggestion paper and a short instruction. I also rule out the possibility that the increased effort is the result of operators working

tation of ideas, suggesting that simply the act of exerting one's voice as part of the suggestion process makes workers more satisfied.[41] Arguably, addressing workers' aspirations is particularly important in unlocking bottom-up information flows across hierarchies for firms in lower-income countries. This offers large potential for further research.

5.1.2 Perceived increase in psychological safety-related constraints

I continue by analysing the effects of the huddle intervention (see column 4 in Table 3). Why does the huddle intervention decrease the number of suggestions on supervisors' behaviors and skills? There are positive and negative channels related to social costs through which the intervention could impact the content of suggestions. On the positive side ("real satisfaction"), the training program may have improved supervisors' skills. Huddles may have improved workers' perception of their supervisors. Hence, such changes could have led to fewer suggestions about supervisors' skills. I measure real satisfaction through a perceived leadership efficacy index, which contains survey questions about supervisors' abilities (supervisor ability index) and how much workers feel cared for (caring culture index). On the negative side ("unspoken dissatisfaction"), huddles may lead the voice climate to deteriorate. The huddles may have steered workers away from broaching topics they would otherwise raise. Supervisors may have manipulated workers to submit suggestions that did not portray the supervisors in a negative light to protect their reputations among managers – a development that would be in line with Friebel and Raith's (2004) theory on skip-level communication.[42] Unspoken dissatisfaction is

toward a mission they care for (Dellavigna et al., 2018; Tonin and Vlassopoulos, 2010, 2015). For this mechanism to be at work, the lunch intervention would have had to have led to changes in workers' perceptions about how seriously the firm takes the suggestion box. I argue that this is not the case. First, if this had been the case, I would have expected to see spillover effects on workers who were in the control group and had friends in the lunch group. Yet, those workers in the control group submit a similar number of suggestions as workers in the control group who had no friends who were part of the lunch sessions. Second, even after the first announcement of awards, which also signals the seriousness of the suggestion box, the behavior of workers in the control group did not change. Finally, I find no evidence of learning effects: the quantity, quality and variety of suggestions do not improve over time (see Appendix Table K.24).

[41]This is supported by a finding that workers in the lunch group who make an above-median number of submissions appear to be more satisfied than those who did not – a finding that does not stem from any greater numbers of experiences or interactions (see columns 1 and 2 in Appendix Table K.33).

[42]Friebel and Raith (2004) suggest in their theoretical model that skip-level communication (i.e., communication between subordinates and top managers) may induce middle managers to deliberately recruit unproductive subordinates, or to abuse their authority in other ways, to protect themselves. Middle managers can have strategic concerns when communication between their subordinates and superiors is too open (Baird, 1977; Lillico, 1972).

measured by a perceived voice climate index, which captures the extent to which workers feel comfortable to share their opinions, and contains survey questions about the ability to be yourself at work, fear of doing something wrong and admitting an error, or degree of trust between workers and managers (detailed overview of all variables in Appendix Figure G.10).

As shown in Table 3, I find no evidence to support the first positive channel of "real satisfaction." Workers do not perceive any changes in abilities of their own supervisor or of a typical supervisor in the factory. The unchanged mental health index (shown in column 9 of Table 2) provides further evidence for this finding. In particular, misbehavior of supervisors – the subject raised by the majority of suggestions on supervisor skills – does not seem to have changed, as indicated by the insignificant estimate on workplace harassment (column 14 of Appendix Table J.19). Workplace harassment is measured by the frequency that uncomfortable situations arise at work, and by the frequency that abusive language and shouting from supervisors occurs. Furthermore, the self-assessment of supervisors who attended the training program did not change (see last row and column 10 of Table 3). By contrast, I find evidence to support the second negative channel of "unspoken dissatisfaction," as shown by the significant deterioration of the perceived voice climate index of 0.2 (see Table 3). Workers exposed to line huddles from trained supervisors seem to have more fear about the consequences of speaking up, and they seem to feel less safe to state what they think. Interestingly, in the endline survey 30% of all supervisors openly admit to having concerns about workers using the suggestion box to complain about them.[43] Qualitative evidence from discussions with managers also supports these findings. Managers were approached by worried supervisors who feared negative assessments would surface through workers' suggestions.

Overall, this evidence makes me confident in concluding that the huddle treatment comes at the cost of steering workers away from topics they might otherwise raise – such as their supervisors' skills. Middle managers (i.e., supervisors) simultaneously act as principals (to workers) and agents (to higher-level managers). They seem to be reluctant to foster the free, upward information flows from workers to higher-level managers –

[43]This percentage is twice as high for untrained supervisors than for trained supervisors, but the difference is not significant.

largely because these flows might involve negative messages about them.

5.2 Heterogeneous treatment effects

5.2.1 Effects by individual disposition of workers

In this section, I show that the lunch treatment successfully decreases incentive-related constraints ("Why should I bother?") as outlined in Section 5.1.1 for both workers with a below- and above-median experience index, but does not help to overcome psychological safety-related constraints. I further provide evidence of a peer-mentor effect, suggesting that communication with friends with greater levels of experience might help to overcome psychological safety-related constraints for workers with lower levels of experience.

Experience of workers. I first look at heterogeneous effects by workers' baseline experience index, which reflects a worker's maturity and skill-set. It is measured by an Anderson (2008) index of the number of factories that have employed the worker, and the worker's grade, tenure in the factory, tenure in the industry, and educational level. Table 4 shows two main findings. First, workers with a above-median experience index do not submit more or better suggestions than workers with a below-median experience index when they are not assigned to the lunch treatment. This is shown by the insignificant coefficients on "High-exp worker" (in columns 1 and 4).[44] Second, the lunch treatment is effective for workers with a below-median experience index, but even more effective for workers with an above-median experience index, as indicated by the significantly positive coefficients on the interaction term "Lunch * High-exp worker" in columns 1 and 4. Within the lunch group, workers with an above-median level of experience submit on average 0.13 more suggestions per round, and the suggestions are of 8% higher quality than those that have below-median levels of work experience. These findings support two hypotheses: (i) Workers with high levels of experience possess the knowledge and confidence to submit

[44] I also explore heterogeneous effects by the number of sewing processes a worker can perform, which likewise reflects a worker's skill set. The factory prefers and gives priority to workers who can perform multiple operations. Other than in Table 4, Appendix Table K.28 shows that workers in the control group, who can perform an above-median number of sewing processes, do indeed submit more suggestions. However, the coefficient ("High-skilled worker") of 0.02 is very small in size.

suggestions (i.e. high levels of psychological safety), but do not bother to submit in the absence of the lunch treatment ("futility of raising your voice"), and (ii) Workers with low levels of experience do not possess the same level of psychological safety as workers with higher levels of experience, even in the presence of the lunch treatment. For example, workers with higher levels of experience may possess better abilities to develop suggestions because they may have more insights about possible changes or lower levels of fear about consequences of bad suggestions. There are no heterogeneous effects by experience for the huddle treatment (see columns 2, 3, 5 and 6).

I also examine heterogeneous effects by workers' own gender. In the sample, 25% of the workers are male, and 75% are female. Because 45% of women and 65% of men have above-median experience indices, I look at heterogeneous treatment effects by gender interacted with experience; I do not find any compelling evidence for gender differences (see columns 2 and 5 of Appendix Table K.29). While suggestions from females with lower levels of experience seem to be of significantly lower quality than those from males with lower levels of experience (see column 5), the coefficient is very small in size. This suggests that experience rather than gender drives the results.

Peer-mentor effect. A worker's own experience and the experience of a worker's peers may be either complements or substitutes. I define a worker's peers as the group of workers who are assigned to be seated opposite, left, right, diagonally left or diagonally right of that worker at each lunch round. Thus, the following analyses only include workers in the lunch group. (Those not in the lunch group do not have any assigned seating partners.) The experience of peers is defined as the percentage of seating partners who have above-median levels of experience. It is defined in percentage terms because some workers have fewer seating partners than others, for example, because some seats are at edge of a table. I generate a median of the percentages, and use an indicator variable equal to one if the percentage is above the median. I find that the experience of peers matters, but there are no substitution or complementarity effects between workers' own and their peers' experience (see Table 6). Both experiences seem to matter individually, although a worker's own experience seem to be a more important driver – as evidenced by the larger

coefficient on own experience compared to the one on peers' experience (see Table 6).

This result raises the question whether there are any circumstances when the workers' own and their peers' experience are actually complements. Particularly in low-trust environments such as in Bangladesh, trustworthiness among discussion partners is considered to be a crucial determinant of exchanging and complementing information. Workers can feel more comfortable and confident (i.e. higher levels of psychological safety) about engaging in discussions with trustworthy partners, such as friends.[45] The effect of friendship is identified because half of the workers in the lunch group were randomly assigned to be seated opposite a friend. A friend is defined as the co-worker a worker talks to most often about work-related issues, and a co-worker who can be considered as a trustworthy partner. I define experience of peers as an indicator equal to one if the person seated *opposite* a worker has an above-median experience index.

I find evidence of a peer-mentor effect. Less-experienced workers benefit from being seated opposite a more-experienced friend during lunches. For example, among workers with low levels of experience, the quality of suggestions increases by 14% (0.36 from a control mean of 2.59) as the result of being seated opposite a highly experienced partner; moreover, the quality of suggestions rises by an additional 24 percentage points if that highly experienced partner is also a friend. There is no such effect for highly experienced workers. For these workers, no differences in the quality of suggestions emerges according to the seating arrangements; it does not matter whether such workers are seated opposite a partner with a low or high level of experience; nor does it matter whether that partner is also a friend (see Appendix Table K.30). This suggests that specific seating arrangements are only effective for certain workers and in certain combinations; only those with low levels of experience benefit – and only when they are seated opposite a highly experienced worker, with the effects more pronounced if that highly experienced worker is also a friend. These patterns are illustrated in Figure 5; corresponding regression results can be found in the Appendix Table K.30.[46] Why is this the case? Workers with low levels of experience

[45]Hahn et al. (2019) find that the individual performance of lower-ability female children in Bangladesh improves once they study with peer friends.

[46]I further examine whether being seated opposite a new, random partner after two lunch sessions stimulates the generation of ideas, and whether these effects are different depending on whether a worker was previously seated opposite a friend or non-friend. I am able to examine this as a result of the built-in design, in which the seating partners of one-third of the workers in the friend and non-friend group are

might have more trustworthy interactions with highly experienced friends, and they might feel more inspired and more confident about submit their ideas in discussion with more senior friends. It remains to be analyzed whether workers are submitting the same ideas as their friends, or whether they just get help formulating their own ideas.

Finally, I examine heterogeneous effects by the gender composition during lunches. I calculate the percentage of male peers seated around a worker by dividing the number of men as seating partners by the total number of seating partners. I then construct an indicator variable equal to one for percentages above the median. I do not find any effects of being seated around an above-median percentage of male peers, regardless of the worker's own gender (see column 2 and 5 of Appendix Table K.32). This reinforces the broader findings that experience drive the results.

5.2.2 Effects by individual disposition of supervisors

I unpack heterogeneous effects of the huddle treatment by supervisors' individual characteristics. There might be differences in the way supervisors conduct huddles, or in the way workers perceive huddles. For example, instructions from more senior supervisors might be clearer, and come across with higher legitimacy; hence, certain supervisors may be better than others in accelerating more efficient peer-to-peer interactions and decreasing psychological safety-related constraints.

I restrict regressions to workers assigned to the lunch group only[47] and I look at heterogeneous effects by supervisors' average experience, gender, baseline perceived efficacy per line, and line performance. The experience index of supervisors is similarly constructed as the index for workers, but does not include grades (because grades are the same for all supervisors). The gender variable measures the percentage of female supervisors per line. The perceived efficacy index is an Anderson (2008) index of baseline survey questions from supervisors about their perceived likelihood of awards on their line, and perceived motivation, concentration, and ability of workers to develop suggestions.

re-randomized after two lunches. I do not find any significant effects. If anything, being seated opposite a new partner is worse for a worker who has previously been seated opposite a friend, as shown by the negative but insignificant coefficients in columns 2 and 4 of Appendix K.31.

[47]Recall that in the absence of the lunch intervention, the huddle treatment had no effects on its own (see Section 4.1).

On treatment lines with more than one supervisor at baseline, I take line averages. The line performance is an Anderson (2008) index of alteration rates, absenteeism rates, and efficiency at baseline.

I find heterogeneous treatment effects by the baseline perceived efficacy (see columns 2 and 6 of Table 5). Line huddles are more effective when they are led by supervisors who express stronger belief in their workers' ability to win, and who assess their workers in a more positive light (in terms of motivation, concentration, and ability to develop suggestions). The result suggests that supervisors have a fairly good sense of how their workers will perform. Furthermore, huddles led by supervisors who have the same gender as workers generate both a higher quantity of ideas and ideas of higher quality (see columns 3 and 7 of Table 5). This result implies that huddles led by female supervisors might be more effective for female workers because the events decreased their social costs. In such a situation, female workers might feel more secure to submit ideas, particularly in an environment characterized by class differences and low levels of trust. The supervisor's experience or line performance does not seem to make a difference.

6 Conclusion

In this paper, I examine whether management practices can encourage bottom-up innovation, leading to beneficial information flows from production workers on lower hierarchical levels to higher-level management with the power to implement ideas. To explore the issue, I set up a suggestion system in a Bangladeshi garment factory, in which the best three ideas from each production floor received a financial bonus and internal recognition. I conducted a randomized controlled trial with two cross-randomized managerial interventions that targeted reducing the costs of contributing suggestions. I examine whether the interventions reduced the effort and time involved, the related social costs, and the combination of both costs. To target costs of effort and time, I provided dedicated time for randomly selected workers to engage in discussions about ideas during free, peer-to-peer lunch sessions structured by external facilitators ("lunch treatment"). To target social costs, I randomly assigned supervisors to attend a two-hour training

program on worker communication methods that were intended to help them encourage workers to give suggestions in monthly three-minute line huddles ("huddle treatment"). To the best of my knowledge, this is the first study to provide empirical evidence from firms in low-income countries about the effects of voice-enhancing managerial practices to address impediments that block upward flows of ideas from production workers to top managers.

I show that structured, peer-to-peer lunches foster the flow of ideas from the bottom up, and, simultaneously, improves worker well-being. While I do not find any short-term effects on firm performance, I find that the lunch treatment significantly improves workers' job satisfaction. In the lunch treatment, a tradeoff between quality and quantity emerged between the treatment and control workers. Workers assigned to the lunch treatment submitted significantly more suggestions than workers who were not involved in the sessions (1,809 vs 48 suggestions over the intervention period of three months). However, workers in the lunch sessions submitted suggestions of lower quality (20% lower quality) than the quality of the suggestions made by those who did not participate in the lunches. This suggests positive self-selection of high-quality submissions from the workers who were in the control group (and thus were not participants in the structured lunch sessions). I provide three explanations that argue for trading off average quality for quantity; that is, I advocate pursuing more suggestions of varying quality from the lunch group rather than fewer suggestions of high quality from the others. The lunch treatment has three advantages. (i) It increases the likelihood of receiving winning suggestions. (ii) It enables the firm to get a feeling for what workers think and care about through the larger breadth and diversity of submitted topics. And, (iii) it makes workers better off regardless of the quality of submitted suggestions (which could eventually lead to reduced turnover in the long-run). By contrast, the huddle treatment that involved meetings led by middle managers did not have any effects on the quantity, quality or variety of suggestions, or on firm performance and worker well-being. In fact, the involvement of middle managers came at a price; their participations led workers to submit fewer suggestions on supervisors' skills and behaviors.

Analysis of the data from surveys of workers reveals the underlying drivers of the

two interventions. The efficacy of the lunch intervention can be explained by a perceived decrease in incentive-related constraints despite suggestions not being implemented. The lower number of suggestions on supervisor skills from workers assigned to the huddle treatment can be explained by a perceived increase in psychological safety-related constraints, i.e. a deterioration in the voice climate, in line with Friebel and Raith's theory on skip-level communication (Friebel and Raith, 2004). Further, I find heterogeneous treatment effects by workers' experience, and evidence of a peer-mentor effect; less-experienced workers benefit from being seated opposite a more-experienced friend during lunches.

My findings raise several important directions for future research. First, this research highlights impediments to upward flows of ideas. Having found that aspirations and confidence are main drivers for idea generation, more targeted follow-up interventions could be conducted which directly address these channels. For example, future research could look at the effects of role models in changing workers' aspirations (e.g., by varying the characteristics of the winners whose ideas and names are announced and displayed). Second, future research could more fully investigate long-run ramifications of such interventions on firms – whether effects on job satisfaction are persistent and eventually translate into reductions of turnover rates, and whether successful implementation of ideas affect long-run productivity. This was a short-run experiment only, in which the firm did not make any major changes across the entire factory based on submitted suggestions. Third, a critical question is how the involvement of middle managers can be made more effective (e.g., through non-monetary or monetary incentive structures). The setting of this study did not allow for changes in the monetary incentives for supervisors. However, discussions with industry experts suggest that this might offer a lever for a more successful involvement of supervisors. Overall, this paper has shown that it is indeed possible and effective to improve upward flows of ideas of workers, who have never before been asked for their opinions on business practices. More research is needed on further targeted managerial interventions to strengthen a culture of bottom-up innovation.

7 Figures

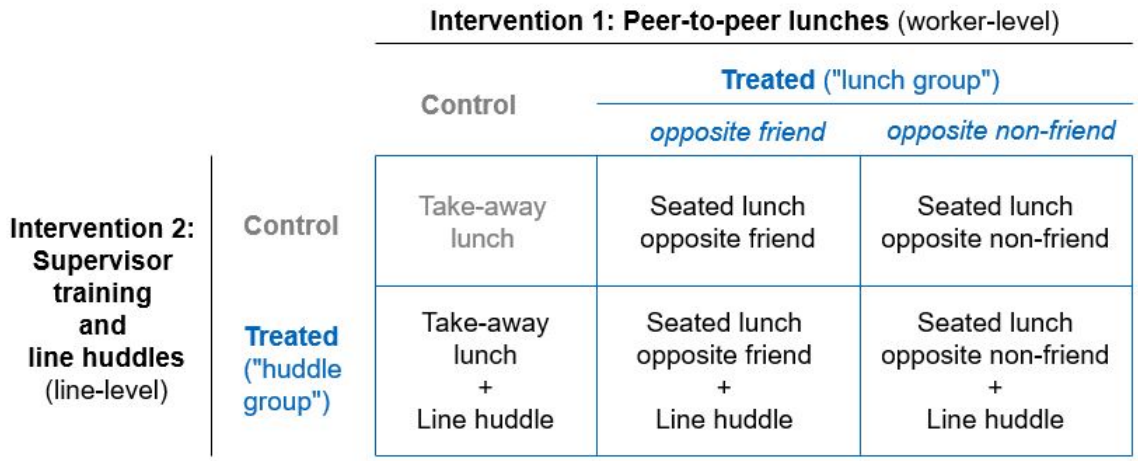


Figure 1: Overview of 2x3 design

Notes: Figure 1 shows the cross-randomised experimental design. Workers are directly randomly assigned to the first managerial intervention, structured peer-to-peer-lunches, either to be seated opposite a friend or a non-friend. Depending on which production line they work at, they are indirectly assigned to the second managerial intervention, line huddles. Thus, workers can be clustered into on of the six buckets in the 2x3 matrix.

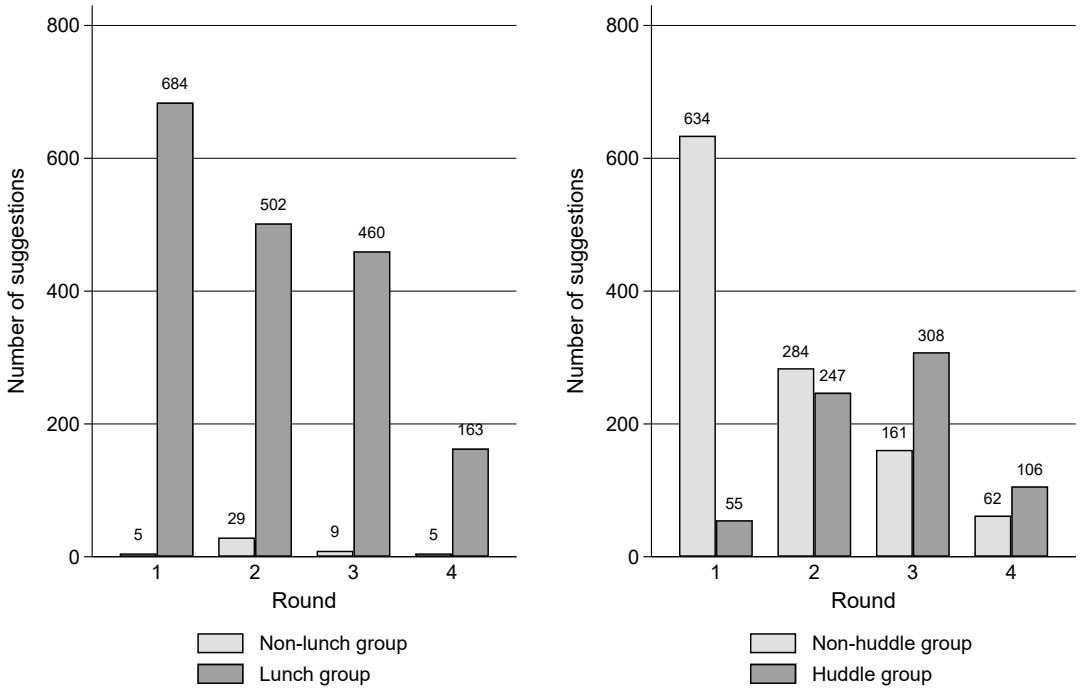


Figure 2: Total number of suggestions per round

Notes: The left graph of Figure 2 shows the total number of suggestions submitted each round by workers assigned to the lunch treatment and control, whereas the right graph shows the same for the huddle treatment and control workers. If a submitted suggestion paper contains several different suggestions, the paper is split by suggestion. In round four, only a quarter of the lunch group was treated (due to Covid-19), leading to a sharp decrease in submissions. Please note that the number of rounds does not always correspond to the actual number of lunches in case of non-compliance, e.g. if a worker was ill/absent, and that it never corresponds to the number of huddles. Few workers received a huddle before the end of Round 1. Overall, the total number of suggestions from the lunch group over the course of the intervention period is 1,809 as compared to 48 from the non-lunch group. The total number of suggestions from the huddle group after the 14th of January, which is the date by which the first huddle took place (at the end of round one), is 716 across all rounds as compared to 507 from the non-huddle group.

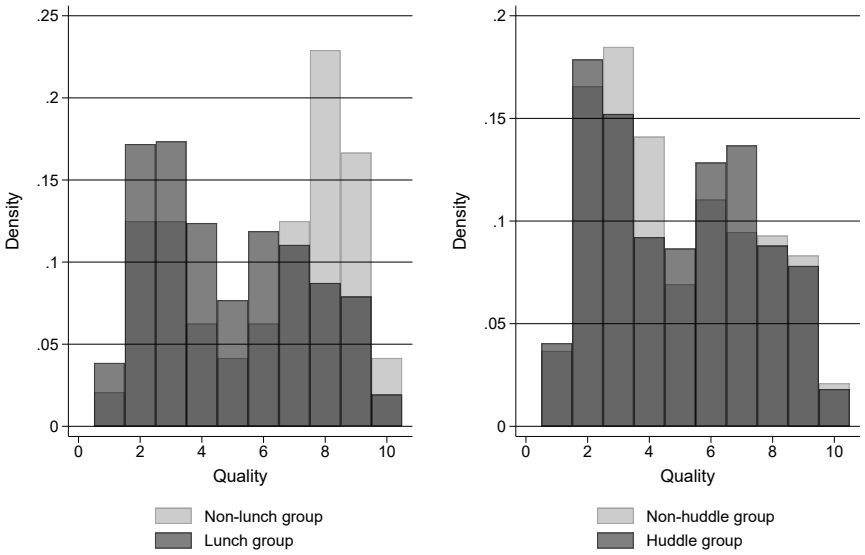


Figure 3: Distribution of quality of submitted suggestions

Notes: Figure 3 shows the distribution of the quality of submitted suggestions by lunch treatment and control group on the left side and by huddle treatment and control group on the right side. Suggestions classified as incoherent receive a quality index rating of 1, general observations/no problems 2, a poorly communicated problem 3, a well communicated problem 4, a poorly communicated idea 5, a well-communicated idea 6. Ideas are then further evaluated based on five criteria on a scale from one to five (implementation level, impact on worker well-being, impact on efficiency, implementation difficulty, and financial resources), and one criterion on a scale from one to two (level of effects). The sum is clustered into five ranges (1-19, 20, 21-22, 23, and 24-27 points). Points are clustered such that each of the five ranges has a similar number of observations. Those ranges are added to the quality index for poorly or well communicated ideas, such that they range from five to nine or six to ten points, respectively.

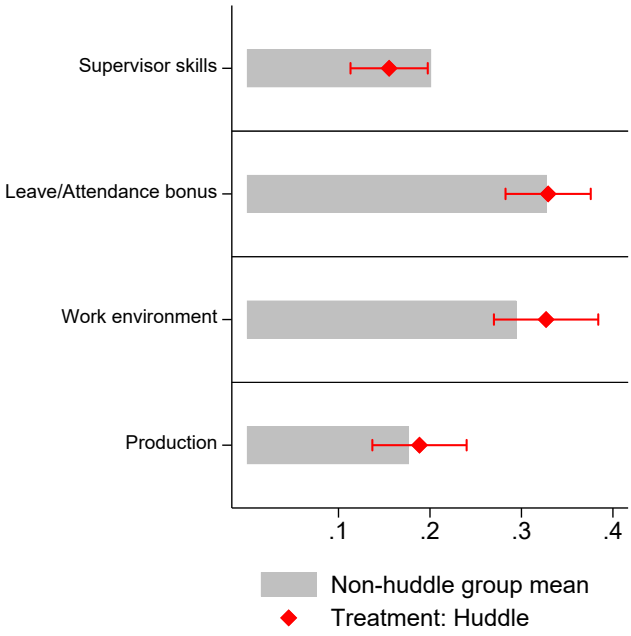


Figure 4: Effects on content of suggestions by treatment groups

Notes: Figure 4 shows the ITT effects of four individual regressions of the treatment indicator on the content variable. The content variable is an indicator equal to one if the submitted suggestions is classified as the respective category. Suggestions are only classified if they are problems or ideas (type three or six). Regressions in the lower graph exclude all suggestions prior to the huddle start date, i.e. 14th of January 2020. 95% confidence intervals are displayed.

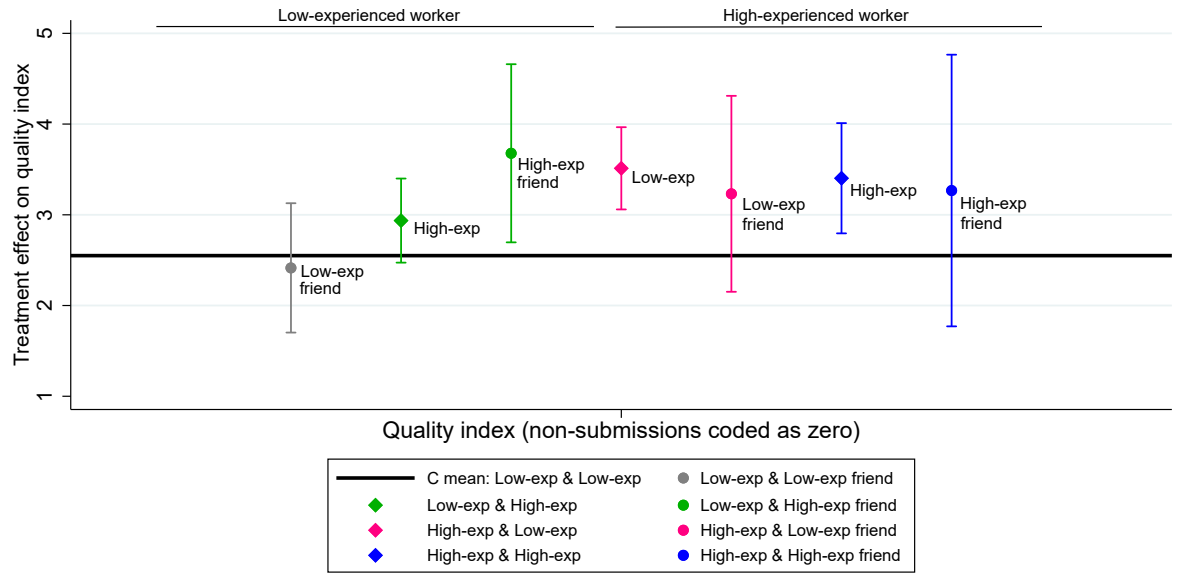


Figure 5: Illustration of heterogeneous treatment effects by experience and friendship

Notes: Figure 5 illustrates the treatment effects of the 4th column of Table K.30. The labels at the bottom of the figure describe the own experience of the worker first, and the experience of the seating partner and friendship second. For clarity, I also display additional headlines and markings inside the figure. The headlines indicate the workers' own experience. The first three bars show effects for low-experienced workers, whereas the last four bars show effects for high-experienced workers. The markings indicate the seating partners experience, and whether the seating partner is a friend. Bars in grey and pink are low-experienced seating partners, and bars in green and blue high-experienced partners. Circles indicate that the partner is a friend, whereas diamonds indicate that the partner is a non-friend. The horizontal line at 2.59 shows the mean of the control group, which are low-experienced workers, sitting opposite a low-experienced non-friend. 95% confidence intervals are displayed.

8 Tables

TABLE 1. EFFECTS ON PRIMARY OUTCOME VARIABLES RELATED TO INNOVATION

	Quantity			**Quality**			**Quality (if submitted)**			**Variety**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lunch	0.45***		0.43***	2.20***		2.21***	-1.04**		-1.56**	
	(0.012)		(0.019)	(0.072)		(0.15)	(0.51)		(0.73)	
Huddle		-0.013	-0.017		-0.091	-0.13		-0.081	-0.33	0.066
		(0.017)	(0.011)		(0.11)	(0.083)		(0.20)	(1.14)	(0.052)
LunchXHuddle			-0.012			-0.045			0.29	
			(0.025)			(0.19)			(1.14)	
Observations	6352	5179	5179	6352	5179	5179	1499	1294	1294	544
R^2	0.34	0.08	0.30	0.24	0.06	0.22	0.06	0.08	0.10	0.04
Control Mean	0.02	0.26	0.18	0.10	1.37	0.97	6.60	5.35	5.37	1.13
Round FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
p (L=H)			0.00			0.00			0.17	
p (L=LxH)			0.00			0.00			0.29	
p (H=LxH)			0.87			0.71			0.79	

Notes: In columns 1, 4 and 7 standard errors are clustered at the pair level. In all other columns, standard errors are clustered at the line level, and observations prior to the first huddle (14th of January 2020) are excluded. Columns 7 to 10 only include submitted suggestions. Column 10 further restricts observations to individuals who submitted more than once (which are only some of the lunch workers). The quality index (column 4 to 9) consists of six components, out of which two are criteria to evaluate the usefulness of submitted ideas for workers, namely impact on worker well-being and implementation level, and four are criteria to evaluate the usefulness of ideas for the firm, namely financial resources, impact on efficiency, implementation time and risk, and level of effects (separate regressions shown in Table J.21). * p<0.1, ** p<0.05, *** p<0.01

TABLE 2. EFFECTS ON PRIMARY OUTCOME VARIABLES RELATED TO FIRM PERFORMANCE AND WORKER WELL-BEING

	Workforce index			**Production index**	**Job satisfaction index**			**Mental health index**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lunch	-0.006		-0.000		0.196**		0.301	0.016		0.069
	(0.734)		(0.988)		(0.037)		(0.102)	(0.892)		(0.605)
Huddle		0.004	0.017	-0.162		0.116	0.219		0.171	0.243
		(0.817)	(0.517)	(0.182)		(0.193)	(0.139)		(0.282)	(0.203)
LunchXHuddle			-0.025				-0.229			-0.165
			(0.514)				(0.268)			(0.444)
Observations	119588	91580	91580	459	195	195	195	195	195	195
R^2	0.09	0.09	0.09	0.50	0.20	0.19	0.22	0.13	0.15	0.18
Control Mean	-0.14	-0.10	-0.10	0.15	-0.23	-0.17	-0.14	-0.13	-0.17	-0.12
Clusters	792	785	785	51	182	51	51	182	51	51
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
p (L=H)			0.52				0.54			0.36
p (L=LxH)			0.68				0.15			0.47
p (H=LxH)			0.48				0.18			0.26

Notes: Standard errors are clustered at the pair level in columns 1, 5, and 8, and at the line level otherwise. Time fixed effects are included in regressions on the workforce index (in days) and on the production index (in weeks). Huddle exposure is truncated at one (at least one huddle). The cumulative exposure per worker is used in columns 1, 2 and 3, the average cumulative exposure per line in column 4, and the maximum exposure per worker in columns 5 to 10. Indices are constructed using the Anderson (2008) method. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE 3. PERCEIVED CONSTRAINTS (SURVEY DATA)

	Operators								Supervisors	
	(1)		(2)		(3)				(4)	
	Control μ	Lunch	Control μ	Huddle	Control μ	Lunch	Huddle	Both	Control μ	Huddle
Incentive-related constraints										
Futility Index	-0.04	0.35***	-0.01	0.08	-0.03	0.29	-0.01	0.11	-0.05	0.02
		(0.12)		(0.13)		(0.19)	(0.17)	(0.17)		(0.22)
Time Index	-0.11	0.09	-0.04	0.20	-0.14	0.06	0.15	0.05	-0.02	-0.05
		(0.12)		(0.14)		(0.16)	(0.20)	(0.26)		(0.24)
Psychological safety-related constraints										
Knowledge Index	-0.02	0.04	-0.00	-0.00	0.02	-0.05	-0.09	0.16	-0.06	0.15
		(0.15)		(0.12)		(0.17)	(0.21)	(0.27)		(0.23)
Security Index	0.01	0.01	0.00	-0.11	0.00	-0.08	-0.22	0.21	0.02	-0.07
		(0.10)		(0.12)		(0.18)	(0.16)	(0.21)		(0.18)
Subindex: Voice climate	0.03	0.06	0.03	-0.20*	0.06	-0.07	-0.34**	0.26	0.08	-0.02
		(0.11)		(0.11)		(0.17)	(0.16)	(0.23)		(0.17)
Subindex: Leadership efficacy	-0.00	-0.11	-0.01	0.02	-0.04	-0.07	0.08	-0.10		
		(0.12)		(0.11)		(0.14)	(0.14)	(0.18)		
Number of Observations	200		200		200				97	

Notes: Table 3 shows the perceived constraints indices based on survey questions with operators (columns 1 to 3) and supervisors (column 4). It exhibits individual regressions of the constraints indices on the lunch treatment indicator in column 1, and on the huddle treatment indicator in columns 2 and 4. Column 3 shows a single regression of the constraints indices on the lunch treatment indicator, the huddle treatment indicator, and an interaction term. Dependent variables, i.e. the constraints indices, are shown in bold on the left. The futility index contains survey questions on the perceived likelihood of winning and implementation of ideas. The time index contains survey questions on the frequency of thinking about problems and solutions. The knowledge index contains survey questions on the easiness to think about problems and solutions. The security index comprises of two sub-indices: The voice climate index measures the degree to which workers feel comfortable to raise their opinions, the leadership efficacy index measure the degree to which workers feel cared for (caring culture index) and an assessment of their supervisors' abilities (supervisor abilities index). Detailed index components are shown in Appendix Figure G.10, and individual regressions are shown in Appendix Table K.25 and K.26. Standard errors are clustered at the pair level in the regression in column 1, and at the line level in all other regressions. All regressions include floor fixed effects to control for any time-invariant differences across floors. * p<0.1, ** p<0.05, *** p<0.01

TABLE 4. HETEROGENEOUS TREATMENT EFFECTS BY WORKERS' EXPERIENCE

	Quantity			**Quality**		
	(1)	(2)	(3)	(4)	(5)	(6)
High-exp worker	-0.011	0.080***	-0.0055	-0.058	0.51***	-0.035
	(0.0081)	(0.022)	(0.0080)	(0.053)	(0.13)	(0.062)
Lunch	0.39***		0.33***	1.81***		1.51***
	(0.017)		(0.023)	(0.093)		(0.12)
Lunch * High-exp worker	0.13***		0.12***	0.78***		0.81***
	(0.023)		(0.034)	(0.14)		(0.20)
Huddle		0.0072	-0.0048		0.061	-0.050
		(0.019)	(0.010)		(0.097)	(0.057)
Huddle * High-exp worker		-0.030	-0.013		-0.22	-0.046
		(0.029)	(0.010)		(0.16)	(0.070)
Lunch * Huddle			-0.022			-0.0092
			(0.031)			(0.16)
Lunch * Huddle * High-exp worker			0.027			-0.067
			(0.042)			(0.23)
Observations	6352	4955	4955	6352	4955	4955
R^2	0.35	0.07	0.29	0.25	0.06	0.21
Control Mean	0.01	0.18	0.02	0.06	0.85	0.10
Round FE	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES
Lunch=Lunch*High-exp worker	0.00		0.00	0.00		0.01
Huddle=Huddle*High-exp worker		0.41	0.63		0.22	0.97

Notes: Table 4 shows heterogeneous treatment effects by workers' experience. Experience is measured by an Anderson (2008) index of the number of factories worked at, grade, tenure in the factory, tenure in the industry and the educational level. "High-exp worker" means that the respective worker has an above-median experience index. Columns 2, 3, 5, and 6 exclude all observations prior to the start date of huddles, i.e. the 14th of January 2020. A the bottom, I show p-values from comparing each treatment and treatment interaction against each other. Regressions include round and floor fixed effects and stratification variables. Standard errors are clustered at the pair level in columns 1 and 4, and at the line level in columns 2, 3, 5 and 6. Standard errors are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE 5. HETEROGENEOUS TREATMENT EFFECTS BY SUPERVISORS' DISPOSITION

	Quantity				**Quality**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Huddle	0.017	-0.039*	-0.035	0.0022	0.063	-0.23*	-0.27**	0.025
	(0.023)	(0.021)	(0.024)	(0.023)	(0.12)	(0.12)	(0.13)	(0.14)
Huddle * Above-median experienced supervisor	-0.058				-0.28			
	(0.039)				(0.22)			
Huddle * Same gender as supervisor		0.061				0.34		
		(0.043)				(0.21)		
Huddle * Above-median perceived efficacy of supervisors			0.051				0.41*	
			(0.041)				(0.23)	
Huddle * Above-median line performance				-0.031				-0.21
				(0.031)				(0.18)
Observations	4748	3755	4748	4748	4748	3755	4748	4748
Control Mean	0.19	0.22	0.22	0.18	0.94	1.13	1.11	0.91
Round FE	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES
p-value (Huddle=Huddle*Interaction)	0.20	0.07	0.17	0.50	0.27	0.04	0.05	0.43

Notes: Table 5 shows heterogeneous treatment effects of the second managerial intervention by supervisors experience, gender and baseline perceived efficacy. Regressions only include workers assigned to the first managerial intervention, i.e. the lunch treatment. The experience index is an Anderson (2008) of the number of factories worked at, tenure in the factory, tenure in the industry and the educational level. I use averages per line if more than one supervisor is assigned per line. The same gender as supervisor is an indicator equal to one if all supervisors per line have the same gender as the worker. The perceived efficacy index measures supervisors' perceptions about the likelihood of a prize on their line and the likelihood of the implementation of winning suggestions by managers. It also includes an assessment of how motivated and concentrated workers are on their line, as well as how able they are to develop suggestions. Line performance is an Anderson (2008) index of efficiency, alteration rate, and absenteeism rate at baseline. Regressions include floor fixed effects and stratification variables. Standard errors are clustered at the pair level and are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE 6. INTERACTION EFFECTS OF WORKERS' OWN AND THEIR PEERS' EXPERIENCE

	Quantity		**Quality**	
	(1)	(2)	(3)	(4)
High-exp peers	0.047**	0.075**	0.27*	0.28
	(0.023)	(0.036)	(0.15)	(0.21)
High-exp worker		0.16***		0.86***
		(0.035)		(0.21)
High-exp worker * High-exp peers		-0.068		-0.13
		(0.046)		(0.28)
Observations	2140	2140	2140	2140
R^2	0.03	0.05	0.06	0.08
Control Mean	0.65	0.57	3.11	2.67
Round FE	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES
Stratification var.	NO	NO	NO	NO
High-exp peers=High-exp worker		0.06		0.36

Notes: Table 6 shows the effect of workers' own and their peers' experience on the quantity and quality of suggestions. Experience is measured by an Anderson (2008) index of the number of factories worked at, grade, tenure in the factory, tenure in the industry and the educational level. "High-exp worker" means that the respective worker has an above-median experience index. "High-exp peers" means that the respective worker is seated around an above-median percentage of high-experienced co-workers, in particular seating partners to the left, right, opposite, diagonally left and diagonally right. At the bottom, I show p-values from comparing the effect of being a high-experienced worker and having high-experienced peers. Regressions exclude observations from workers not assigned to the lunch treatment. They include round and floor fixed effects. Standard errors are clustered at the pair level and shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

References

Adhvaryu, Achyuta, Teresa Molina, and Anant Nyshadham, "Expectations, Wage Hikes, and Worker Voice: Evidence from a Field Experiment," *NBER Working Paper 25866*, 2019.

Anderson, Michael L, "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," 2008, *103* (484), 1481–1495.

Aoki, Masahiko, "Toward an Economic Model of the Japanese Firm," *Journal of Economic Literature*, 1990, *28* (1), 1–27.

Arrow, Kenneth J, *The Limits of Organization*, Norton, 1974.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit Khandelwal, and Eric Verhoogen, "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan," *Quarterly Journal of Economics*, 2017, *123* (3), 1101–1164.

Baird, John E, *The Dynamics of Organizational Communication*, Harper & Row, 1977.

Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun, "CEO Behavior and Firm Performance," *Journal of Political Economy*, 2020, *128* (4).

Battiston, Diego and Tom Kirchmaier, "Face-to-Face Communication in Organisations," *Working Paper*, 2018, pp. 1–62.

BBC, "Eid al-Adha: Why are there two Eids?," May 2019.

Benjamini, Yoav, Abba M Krieger, and Daniel Yekutieli, "Adaptive linear step-up procedures that control the false discovery rate," *Biometrika*, 2006, *93* (3), 491–507.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts, "Do Management Interventions Last? Evidence From India," *World Bank Policy Research Working Paper no.8339*, 2018.

_ , Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, "Does managment matter? Evidence from India," *The Quarterly Journal of Economics*, 2013, *128* (1), 1–51.

Bolton, Patrick and Mathias Dewatripont, "The Firm as a Communication Network," *The Quarterly Journal of Economics*, 1994, *109* (4), 809–839.

Boudreau, Laura, "Multinational enforcement of labor law: Experimental evidence from Bangladesh's apparel sector," *Working Paper*, 2020.

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar, "The Impact of Consulting Services on Small and Medium Enterprises : Evidence from a Randomized Trial in Mexico Antoinette Schoar," *Journal of Political Economy*, 2018, *126* (2).

Burris, Ethan R, Evan A Bruno, James R Detert, and Andrea Caldwell, "From Barriers to Drivers : Understanding Different Approaches to Getting Systematic Upward Voice in Organizations," *The Oxford Handbook of Organizational Citizenship Behavior*, 2018.

Bursztyn, Leonardo and Robert Jensen, "Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure," *Annual Review of Economics*, 2017.

Cai, Jing and Shing yi Wang, "Improving Management through Worker Evaluations: Evidence from Auto Manufacturing," *NBER Working Paper 27680*, 2020, pp. 1–54.

Coffman, Katherine Baldiga, "Evidence on self-stereotyping and the contribution of ideas," *The Quarterly Journal of Economics*, 2014, pp. 1625–1660.

Cottini, Elena, Takao Kato, and Niels Westergaard-nielsen, "Adverse workplace conditions , high-involvement work practices and labor turnover: Evidence from Danish linked employer-employee data," *Labour Economics*, 2011, *18* (6), 872–880.

Dellavigna, Stefano, John A List, Ulrike Malmendier, and Rao, "Estimating Social Preferences and Gift Exchange at Work," *NBER Working Paper No. 22043*, 2016.

_ , U C Berkeley, and Devin Pope, "What Motivates Effort? Evidence and Expert Forecasts," *Review of Economic Studies*, 2018, (85), 1029–1069.

Edmondson, Amy, "Psychological Safety and Learning Behavior in Work Teams," *Administrative Science Quarterly*, 1999, *44*, 350–383.

_ and Zhike Lei, "Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct," *The Annual Review of Organizational Psychology and Organizational Behavior*, 2014, *1*, 23–43.

Friebel, Guido and Michael Raith, "Abuse of Authority and Hierarchical Communication," *The RAND Journal of Economics*, 2004, *35* (2), 224–244.

_ , Matthias Heinz, and Nick Zubanov, "Middle Managers, Personnel Turnover and Performance: A Long-term Field Experiment in a Retail Chain," *Management Science*, 2021.

Gant, Jon, Casey Ichniowski, and Kathryn Shaw, "Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations," 2002, *11* (2), 289–328.

Geanakoplos, John and Paul Milgrom, "A Theory of Hierarchies Based on Limited Managerial Attention," *Journal of the Japanese and International Economies*, 1991, *5*, 205–225.

Gibbs, Michael, Susanne Neckermann, and Christoph Siemroth, "A field experiment in motivating employee ideas," *The Review of Economics and Statistics*, 2017, *99* (4), 577–590.

Gneezy, Uri and John A List, "Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments," *Econometrica*, 2006, *74* (5), 1365–1384.

Grant, Robert M, "Toward a Knowledge-Based Theory of the Firm," *Strategic Management Journal*, 1996, *17*, 109–122.

Hahn, Youjin, Asadul Islam, Eleonora Patacchini, and Yves Zenou, "Friendship and Female Education: Evidence from a Field Experiment in Bangladeshi Primary Schools," *The Economic Journal*, 2019, *130* (627), 740–764.

Harju, Jarkko, Simon Jäger, and Benjamin Schoefer, "Voice at Work," *Centre for Economic Policy Research*, 2021.

Hayek, Author F A, "The Use of Knowledge in Society," *The American Economic Review*, 1945, *35* (4), 519–530.

Higuchi, Yuki, Hoang Vu Nam, and Tetsushi Sonobe, "Management skill, entrepreneurial motivation, and enterprise survival: Evidence from randomized experiments and repeated surveys in Vietnam," *Working Paper*, 2017.

Hjort, Jonas, "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 2014, pp. 1899–1946.

Hoffman, Mitchell and Steven Tadelis, "People Management Skills , Employee Attrition , and Manager Rewards : An Empirical Analysis," *Journal of Political Economy*, 2020, *129* (1).

Iacovone, Leonardo, William Maloney, and David McKenzie, "Improving Management with Individual and Group-Based Consulting: Results from a Randomized Experiment in Colombia," *Working Paper*, 2018, pp. 1–69.

Lazear, Edward P, Kathryn L Shaw, and Christopher T Stanton, "The Value of Bosses," *Journal of Labor Economics*, 2015, *33* (4).

Lillico, T M, *Managerial Communication* 1972.

McAfee, Preston R and John McMillan, "Organizational Diseconomies of Scale," *Journal of Economics & Management Strategy*, 1995, *4* (3), 399–426.

Melumad, Nahum, Dilip Mookherjee, and Stefan Reichelstein, "A Theory of Responsibility Centers," *Journal of Accounting and Economics*, 1992, *15*, 445–484.

Morrison, Elizabeth W, "Employee Voice and Silence," *The Annual Review of Organizational Psychology and Organizational Behavior*, 2014, *1*, 173–197.

Prat, Andrea, Jacques Crémer, and Luis Garicano, "Language and the Theory of the Firm," *The Quarterly Journal of Economics*, 2007, *122* (1), 373–407.

Qian, Yingyi, "Incentives and Loss of Control in an Optimal Hierarchy," *The Review of Economic Studies*, 1994, *61* (3), 527–544.

Rigtering, J P Coen, G Utz Weitzel, and K Katrin Muehlfeld, "Increasing quantity without compromising quality: How managerial framing affects intrapreneurship," *Journal of Business Venturing*, 2019, *34* (2), 224–241.

Sandvik, Jason J, Richard E Saouma, Nathan T Seegert, and Christopher T Stanton, "Workplace knowledge flows," *The Quarterly Journal of Economics*, 2020, pp. 1–46.

Schwarz, Carlo, "lsemantica: A Stata Command for Text Similarity based on Latent Semantic Analysis," *The Stata Journal*, 2019, *19* (1), 129–142.

Sting, Fabian, Christoph Fuchs, Maik Schlickel, and Oliver Alexy, "How to Overcome the Bias We Have Toward Our Own Ideas," *Harvard Business Review*, 2019.

Tonin, Mirco and Michael Vlassopoulos, "Disentangling the sources of pro-socially motivated effort: A field experiment," *Journal of Public Economics*, 2010, *94* (11-12), 1086–1092.

__ and __ , "Corporate Philanthropy and Productivity: Evidence from an Online Real Effort Experiment," *Management Science*, 2015, *61* (8), 1795–1811.

Weber, Roberto A and Colin F Camerer, "Cultural Conflict and Merger Failure: An Experimental Approach," *Management Science*, 2003, *49* (4), 400–415.

Williams, Martin J and Liah Yecalo-Tecle, "Innovation, Voice, and Hierarchy in the Public Sector: Evidence from Ghana's Civil Service," *Economic Development & Institutions Working Paper Series*, 2019, *WP19/09.*

A Theoretical framework

In the following, I present a general theoretical framework for understanding employee voice, which has been discussed non-formally to date.

A.1 Set-up

Workers on the production floor can exert effort e to submit their ideas about what they would like to improve inside the factory. For simplicity, I assume that workers think about ideas either individually ($N_{ij} = 0$) or with a co-worker j ($N_{ij} = 1$). The co-worker can either be a friend ($G_{ij} = 1$) or a non-friend ($G_{ij} = 0$). This set-up allows us to look at the drivers of idea generation among three different groups of workers:

- Individuals: $N_{ij} = 0$
- Friend pairs: $N_{ij} = 1$ and $G_{ij} = 1$
- Non-friend pairs: $N_{ij} = 1$ and $G_{ij} = 0$

I further assume that all workers have the possibility to win a prize for the best submitted idea. Workers who think about ideas with a co-worker have to share the prize with their sparring partner, if they submit the same idea.[48]

A.2 Utility Function

Individual i derives utility U from submitting ideas. On the one hand, worker i receives utility if she wins the tournament (benefit of winning B), which depends on the probability of winning P. On the other hand, worker i also receives utility from intrinsic motivation independently of the probability of winning (effort incentive M). In fact, the policy of asking workers to generate ideas may have motivated them to exert effort independently of the effect of their effort on the probability of winning a prize. Finally, worker i faces costs of effort C to be deducted from the benefits. Hence, the utility function is specified as below:

$$\begin{aligned} U_i(e_i, e_j, e_{-ij}; \theta_{ij}, \delta_i, N_{ij}) = P(e_i, e_j, e_{-ij}; \theta_{ij}, N_{ij}) \times B(v_f, v_i, v_a; N_{ij}) \\ + M(e_i; \lambda_i, N_{ij}) - C(e_i) \end{aligned} \tag{5}$$

where

- $P(e_i, e_j, e_{-ij}; \theta_{ij}, \delta_i, N_{ij}) \in (0, 1)$ is the **probability of winning a prize** which I define as a function of efforts, i.e. individual i and paired worker's j own efforts and the expected average level of effort of all workers other than the individual and paired worker e_{-ij} (called $\bar{e}$). Due to the competitive tournament structure (only the best ideas win), I assume that the expected effort of everyone else is a substitute on a

[48] In reality, the factory did not award shared prizes, but this was the policy they announced in case the winning idea was submitted by someone else as well.

global level, i.e. higher $\bar{e}$ lowers the probability of winning for individual/paired workers. Hence, $\bar{e}$ enters the probability function negatively. The intensity of the decline of worker i's effort in response to changes in $\bar{e}$ (degree of strategic substitutability) is captured by the **global interaction parameter** δ_i. On a local level within paired workers, I assume that worker j's effort can be a substitute or a complement, i.e. worker j's effort can lower or increase the probability of winning of worker i. In other words, workers can be less productive (substitutes) or more productive (complements) jointly. The change in effort of worker i in response to changes of e_j (degree of strategic substitutability or complementarity) depends on the **local interaction parameter** θ_{ij}. Both interaction parameters are described in more detail in the following:

  - δ_i is defined as the **global interaction parameter** of worker i. For simplicity, I assume $\delta_i = \delta_j$. It captures the intensity to which worker i decreases her own effort given the expected average of everyone else's effort. It depends on the worker's belief about her own probability of winning. If the worker strongly believes that she wins, she will be less inclined to adjust her effort levels to the expected average (small δ_i). If the worker is sceptical of winning (large δ_i), the worker will be inclined to exert less effort the higher the expected average effort (due to the competitive tournament structure).
  - θ_{ij} is defined as the **local interaction parameter** between worker i and paired worker j. It captures the intensity to which worker i increases or decreases her own effort given the effort of worker j. It depends on how well paired workers interact, communicate and coordinate with each other. Factors such as degree of friendship or differences in educational levels might play a role. For example, if worker j works harder and is a good friend of worker i, the latter mights also want to work harder. A possible reason could be more trust among friends leading to better interaction patterns, e.g. they could be less ashamed to screen out less good ideas quicker, to say different perspectives, or to discover "blind spots" in thinking more easily. As another example, if worker j has a higher educational level than worker i leading to less common understanding, worker i might want to free-ride and not exert any effort.

- $B(v_f, v_i, v_a)$ is the **benefit of winning a prize** where

  - v_f: **Extrinsic benefit** of winning a prize, i.e. financial reward
    - α captures the share to which the financial reward is divided among pairs
      - $\alpha = 1$: Ideas submitted by individuals
      - $\alpha = 0.5$: Ideas submitted by pairs. Workers are facing a loss of the extrinsic benefit v_f by 50%, because they have to share the monetary reward with their paired co-worker.
  - v_i: **Intrinsic benefit of individual/joint achievement**, e.g. social approval and recognition (by displaying the winner's name), feeling accomplished about oneself/pair, doing something good for the employer and co-workers, etc.
    - β captures degree of preference for individual/joint consumption. It depends on the friendship between worker and co-worker, but also on differences in education (e.g. a worker might not enjoy joint consumption if the co-worker is not as educated)
      - $\beta > 1$: worker prefers winning jointly/sharing the experience; higher utiliy in case of joint consumption

· $\beta = 1$: indifference between joint and individual consumption

· $\beta < 1$: worker prefers winning alone; lower utility in case of joint consumption

– v_a: **Intrinsic benefit of altruistic behaviour**, e.g. caring about the partner to win and to receive a financial benefit

∗ γ captures the degree of altruism for partner

· $\gamma > 1$: care about partner to win

· $\gamma = 1$: indifference

· $\gamma < 1$: dislike partner to win

- $M(e_i; \lambda_i, N_{ij})$ is the **intrinsic motivation of generating ideas** (regardless of winning), e.g. enjoyment from being listened to and cared for, inner satisfaction from thinking about ideas individually or jointly, joy of doing something good for the employer and co-workers, etc. It is a function of worker i's effort. λ_i captures the size of the intrinsic motivation effect and raises the marginal return to effort, and hence the equilibrium effort.

- $C(e_i)$ is the **cost of effort**, which is a function of individual i's effort to capture differences in individual abilities, i.e. the cost to think about ideas, to draft them in a way managers can understand and reward and to hand them in (including opportunity costs of time). I assume regularity conditions to hold $C'() > 0, C''() > 0$, $lim_{e\to\infty} C'(e) = \infty$ guaranteeing the existence of a unique solution.

In the following, I define the probability and cost functions directly in terms of effort assuming concrete functional forms. I normalize all parameters and effort levels to have a valid probabilty density function. The form of the probabilty function was chosen such that utilitiy is always increasing with the worker's own effort, but can be decreasing or increasing with the co-worker's effort depending on the model parameters. I have selected a simple additive form for the benefit function ($B = \alpha v_f + \beta v_i + \gamma v_a$).
The nested utility function is as follows

$$U = [\frac{1}{3}(1 - \delta_i \bar{e})e_i + \frac{1}{3}N_{ij}((1 - \delta_j \bar{e})e_j + \theta_{ij} e_i e_j)]B + \lambda_i e_i - \frac{e_i^2}{2} \tag{6}$$

with e_i, e_j and $\bar{e}$, δ_i, δ_j, λ_i and $B \in [0, 1]$, and $\theta_{ij} \in [-1, 1]$.

A.3 Optimization

Individual i maximizes utility function U with respect to its own effort e_i:

max U w.r.t. e_i

$$\frac{\partial U}{\partial e_i} = \frac{1}{3}B[(1 - \delta_i \bar{e}) + N_{ij}\theta_{ij} e_j] + \lambda_i - e_i \tag{7}$$

The optimal response functions and how they change depending on parameter values are graphically depicted in the appendix. The second partial derivative with respect to e_j is

$$\frac{\partial U}{\partial e_i \partial e_j} = \frac{1}{3} B N_{ij} \theta_{ij} \tag{8}$$

Whether efforts of worker i and co-worker j are strategic complementarities or strategic substitutes depends on the value of parameter θ:

- $\frac{\partial U}{\partial e_i \partial e_j} < 0$ if $\theta_{ij} < 0$: **Strategic substitutes.** Worker i exerts less effort if worker j exerts more effort, i.e. there is a low level of communication and coordination. Note that worker j could also free-ride and not exert any effort if worker i is working ($\theta = 0$)
- $\frac{\partial U}{\partial e_i \partial e_j} > 0$ if $\theta_{ij} > 0$: **Strategic complements**. Worker i exerts more effort if worker j exerts more effort, i.e. there is a high level communication and coordination among workers.

Assuming the same global interaction parameters ($\delta_i = \delta_j$) and the same size of the intrinsic motivation effect ($\lambda_i = \lambda_j$) per pair, I get the following Nash equilibrium level of effort (proof see Appendix Section A.7):

$$e_i^\star = \frac{B(1 - \delta_i \bar{e}) + 3\lambda_i}{3 - N_{ij} B \theta_{ij}} \tag{9}$$

Using an additive functional form for B yields

$$e_i^\star = \frac{(\alpha v_f + \beta v_i + \gamma v_a)(1 - \delta_i \bar{e}) + 3\lambda_i}{3 - N_{ij}(\alpha v_f + \beta v_i + \gamma v_a)\theta_{ij}} \tag{10}$$

This equilibrium equation indicates that worker i does not exert any effort if $\delta_i \bar{e} = 1$ and $\lambda_i = 0$. The optimal response functions and how the equilibrium efforts change if for example θ or B increase are illustrated in figures A.6.1 and A.6.1 in case of strategic complements, and in Figure A.6.2 in case of strategic substitutes.

For each group the parameters are specified as below.

1. **Individuals** ($N_{ij} = 0$): $e_j = 0$, $\alpha = 1$, $\beta = 1$, $v_a = 0$, $\delta_i = \delta_C$ and $\lambda_i = \lambda_C$
   Individuals are not paired with anyone ($e_j = 0$). They receive the full financial reward v_f ($\alpha = 1$) and they can only receive the intrinsic benefit from individual consumption, not from joint consumption ($\beta = 1$). Also they cannot receive any benefit from altruistic behaviour ($v_a = 0$). The global interaction parameter is δ_C and the intrinsic motivation effect has the size λ_C.
2. **Friend pairs** ($N_{ij} = 1$, $G_{ij} = 1$): $e_j > 0$, $\theta_i = \theta_F$, $\alpha = 0.5$, $\beta = \beta_F$, $\gamma = \gamma_F$, $\delta_i = \delta_F$ and $\lambda_i = \lambda_F$
   Individuals are paired with a friend who exerts effort e_j. The level of communication and coordination among friends is captured by the local interaction term

θ_F. Individuals have to share the prize with their partner ($\alpha = 0.5$). They have consumption preferences β_F and their degree of altruism takes the value γ_F. The global interaction parameter is δ_F and the intrinsic motivation effect has the size λ_F.

3. **Non-friend pairs** ($N_{ij} = 1$, $G_{ij} = 0$): $e_j > 0$, $\theta_i = \theta_{NF}$, $\alpha = 0.5$, $\beta = \beta_{NF}$, $\gamma = \gamma_{NF}$, $\delta_i = \delta_{NF}$ and $\lambda_i = \lambda_{NF}$
Individuals are paired with a non-friend who exerts effort e_j. The level of communication and coordination among non-friends is captured by the local interaction term θ_F. Individuals have to share the prize with their partner ($\alpha = 0.5$). They have consumption preferences β_{NF} and their degree of altruism takes the value γ_{NF}. The global interaction parameter is δ_{NF} and the intrinsic motivation effect has the size λ_{NF}.

I assume that the parameters fulfill the following criteria:

- $\theta_F > \theta_{NF}$: The local interaction parameter is lower for friends than for non-friends, with a lower parameter indicating a higher level of communication and coordination, e.g. due to a higher level of trust.
- $\delta_F < \delta_{NF}$: The global interaction parameter is lower for friend than the non-friend pairs, indicating that the their level of effort is less susceptible to changes in the average level of effort, e.g. because they work well together and think that their probability of winning is high.
- $\beta_F > \beta_{NF}$: The preference for joint consumption for friends is higher than for non-friends, e.g. because it is more rewarding to share the experience of winning with someone a worker cares about.
- $\gamma_F > \gamma_{NF}$: The degree of altruism for a friend is higher than for a non-friend, e.g. because a worker would care more about someone she knows well to win the prize.
- $\lambda_F > \lambda_{NF}$: The size of the intrinsic motivation effect for a friend is higher than for a non-friend, e.g. because thinking about ideas with a friend is much more fun than with a non-friend (or alone).

I can now plug in the parameters to find the optimal level of effort per group.

A.3.1 Optimal effort of individuals

$$e_i^\star = \frac{1}{3} B_C (1 - \delta_C \bar{e}) + \lambda_C = \frac{1}{3} (\alpha v_f + \beta v_i + \gamma v_a)(1 - \delta_C \bar{e}) + \lambda_C \tag{11}$$

A.3.2 Optimal effort of friend pairs

$$e_i^\star = \frac{B_F (1 - \delta_F \bar{e}) + 3\lambda_F}{3 - B_F \theta_F} = \frac{(\alpha_F v_f + \beta_F v_i + \gamma_F v_a)(1 - \delta_F \bar{e}) + 3\lambda_F}{3 - (\alpha_F v_f + \beta_F v_i + \gamma_F v_a)\theta_F} \tag{12}$$

A.3.3 Optimal effort of non-friend pairs

$$e_i^\star = \frac{B_{NF}(1-\delta_{NF}\bar{e}) + 3\lambda_{NF}}{3 - B_{NF}\theta_{NF}} = \frac{(\alpha_{NF}v_f + \beta_{NF}v_i + \gamma_{NF}v_a)(1-\delta_{NF}\bar{e}) + 3\lambda_{NF}}{3 - (\alpha_{NF}v_f + \beta_{NF}v_i + \gamma_{NF}v_a)\theta_{NF}} \tag{13}$$

A.4 Predictions

In this section, I want to test the following two predictions:

1. **Friendship effect** (effect of being paired with a friend): The quality and quantity of suggestions is higher for friend pairs than for non-friend pairs.
2. **Social effect** (effect of being paired with a co-worker): The quality and quantity of suggestions is higher for pairs than for individuals.

These two predictions can be mathematically expressed as:

1. $e_i^\star\big|_{N_{ij}=1,G_{ij}=1} > e_i^\star\big|_{N_{ij}=1,G_{ij}=0}$
2. $e_i^\star\big|_{N_{ij}=1} > e_i^\star\big|_{N_{ij}=0}$

A.4.1 Friendship effect

Inserting optimal efforts into prediction 1) yields

$$\frac{B_F(1-\delta_F) + 3\lambda_F}{B_{NF}(1-\delta_{NF}) + 3\lambda_{NF}} > \frac{3 - B_F\theta_F}{3 - B_{NF}\theta_{NF}} \tag{14}$$

Given that B_F is greater than B_{NF} (as $\beta_{NF} < \beta_F$ and $\gamma_{NF} < \gamma_F$), δ_F is smaller than δ_{NF} and λ_F is greater than λ_{NF}, the LHS is always larger than one. This would also the case even if δ_F were equal to δ_{NF}. Given that $\theta_F > \theta_{NF}$, the RHS is always smaller than one. Hence, prediction 1) holds independent of whether efforts of non-friend pairs are complements (weaker than of friend pairs) or substitutes.

A.4.2 Social effect

As I just showed that the optimal effort of friend pairs always higher than for non-friend pairs, I just have to show that the effort for non-friend pairs is higher than for individuals to proof prediction 2: $e_i^\star\big|_{N_{ij}=1,G_{ij}=1} > e_i^\star\big|_{N_{ij}=1,G_{ij}=0} > e_i^\star\big|_{N_{ij}=0}$

Inserting optimal efforts of the non-friend pairs and of individuals into prediction 2 yields

$$\frac{3}{3 - B_{NF}\theta_{NF}} > \frac{B_C(1-\delta_C) + 3\lambda_C}{B_{NF}(1-\delta_{NF}) + 3\lambda_{NF}} \tag{15}$$

If efforts of non-friend pairs are complements $\theta_{NF} > 0$, the LHS is always greater than one. If efforts of non-friend pairs are substitutes $\theta_{NF} < 0$, the LHS is always smaller than one.

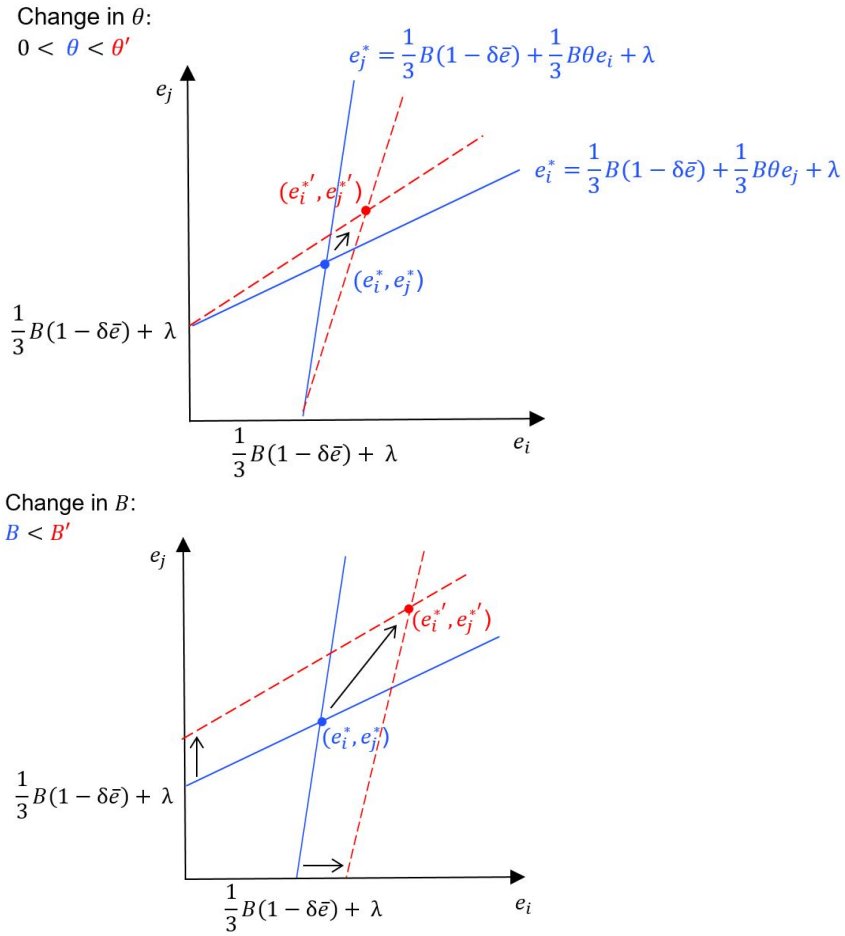
The value of the RHS depends on the size of the benefits (i.e. benefit of winning B and intrinsic motivation effect λ) of individuals relative to non-friend pairs, and the size of the global interaction parameter δ (i.e. how much effort changes in response to changes in the average level of effort): $(v_f + v_i) + \frac{3\lambda_C}{1-\delta_C} \gtrless (\frac{v_f}{2} + \beta_{NF} v_i + \gamma_{NF} v_a) + \frac{3\lambda_{NF}}{1-\delta_{NF}}$ which can be rewritten as $\frac{v_f}{2} \gtrless (\beta_{NF} - 1)v_i + \gamma_{NF} v_a + \frac{3\lambda_{NF}}{1-\delta_{NF}} - \frac{3\lambda_C}{1-\delta_C}$. It shows that the result depends on the size of the loss in financial benefit from sharing the prize relative to the gain in benefits from working with a co-worker (joint consumption/winning, altruistic behaviour and intrinsic motivation) as well as the effort adjustments to the average level of effort. If efforts of non-friend pairs are substitutes (LHS < 1) indicating that the interaction/coordination between worker and co-worker does not work well, the gain in benefit from working with the co-worker must be sufficiently large (e.g. $\beta_{NF} > 1$, $\gamma_{NF} > 1$, and $\lambda_{NF} > \lambda_C$) and/or the global interaction parameter sufficiently small (e.g. $\delta_{NF} < \delta_C$) for the non-friend pairs to exert more effort (so that $1 >$ LHS $= e^*_{NF} \gg$ RHS $= e^*_C$). If efforts are complements and the probability of winning increases in the co-workers effort j (LHS > 1), the gains in benefit from working with a co-worker do not need to be as large for the non-friend pairs to exert more effort (so that LHS $= e^*_{NF} > 1 >$ RHS $= e^*_C$).

A.5 Validity of model assumptions

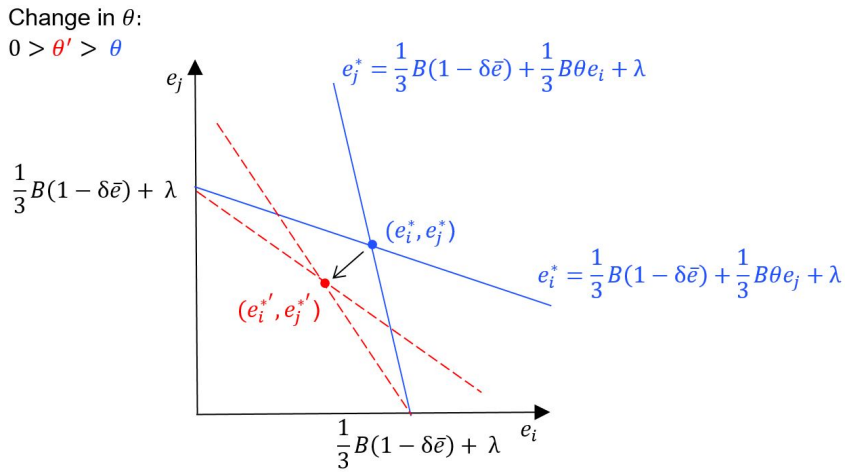
It is important to acknowledge that the average effort of everybody else $\bar{e}$ is really endogenous as the experiment itself changes the average level of idea generation on the production floor. I argue that in this model the endogeneity of $\bar{e}$ does not matter, but rather the assumption that $\bar{e}$ is the same for all groups of workers. This model allows for intragroup comparisons of three groups of workers (friend pairs, non-friend pairs or individuals) on one production floor under a tournament structure. However, I must be careful when extrapolating from this setting to other settings with a different $\bar{e}$.

A.6 Graphical display of strategic complements and substitutes

A.6.1 Strategic complements ($\theta > 0$)



A.6.2 Strategic substitutes ($\theta < 0$)



A.7 Mathematical Appendix: Proof of equation 9

$$\frac{\partial U}{\partial e_i} = \frac{1}{3}B[(1-\delta_i\bar{e}) + N_{ij}\theta_{ij}e_j] + \lambda_i - e_i \tag{16}$$

$$\frac{\partial U}{\partial e_j} = \frac{1}{3}B[(1-\delta_j\bar{e}) + N_{ij}\theta_{ij}e_i] + \lambda_j - e_j \tag{17}$$

Plugging (17) into (16) with $\delta_i = \delta_j$ and $\lambda_i = \lambda_j$ yields

$$\frac{1}{3}B[(1-\delta_i\bar{e}) + N_{ij}\theta_{ij}(\frac{1}{3}B[(1-\delta_i\bar{e}) + N_{ij}\theta_{ij}e_i] + \lambda_i)] + \lambda_i = e_i$$

$$\frac{1}{3}B(1-\delta_i\bar{e})(1+\frac{1}{3}N_{ij}B\theta_{ij}) + \lambda_i(1+\frac{1}{3}N_{ij}B\theta_{ij}) = e_i(1-N_{ij}^2\frac{1}{9}B^2\theta_{ij}^2)$$

$$e_i = \frac{(\frac{1}{3}B(1-\delta_i\bar{e}) + \lambda_i)(1+\frac{1}{3}N_{ij}B\theta_{ij})}{(1-N_{ij}^2\frac{1}{9}B^2\theta_{ij}^2)}$$

$$e_i = \frac{(\frac{1}{3}B(1-\delta_i\bar{e}) + \lambda_i)(1+\frac{1}{3}N_{ij}B\theta_{ij})}{(1+\frac{1}{3}N_{ij}B\theta_{ij})(1-\frac{1}{3}N_{ij}B\theta_{ij})}$$

$$e_i^\star = \frac{B(1-\delta_i\bar{e}) + 3\lambda_i}{3 - N_{ij}B\theta_{ij}}$$

B Determination of line location of workers

The production lines of workers were determined based on daily line movement data from the IE department. However, only 40% of all days per worker are covered during the intervention period. 159 out of 1599 workers are not included at all and some days are missing for individual workers. Data of one unit is completely missing for January and March 2020. 40% of the 159 workers without any information work in that particular unit. Besides, more than one line is recorded for some workers. Due to this incompleteness, the dataset has to be complemented with survey and intervention data. I complement the data with the following approach. First, I use the line code as indicated on suggestion papers for the days at which workers submitted their papers. Second, I 'fill up' line codes with the last available observation assuming that workers continued on the line mentioned last. If the last entry contained several lines, I assume that workers continued on the line which appears next. Mostly, workers were working on two lines just for one day, and continued on the second line going forward. Third, I use line codes from the first available date in February 2020 for all workers from one unit for which the factory did not record any data in January 2020. Fourth, I use line codes from the endline data for missing lines in March 2020. Fifth, I use line codes from the baseline data for the 159 workers without any observations in the line movement data. The distribution of the different data sources to determine each worker's line per day is shown in Figure B.1.

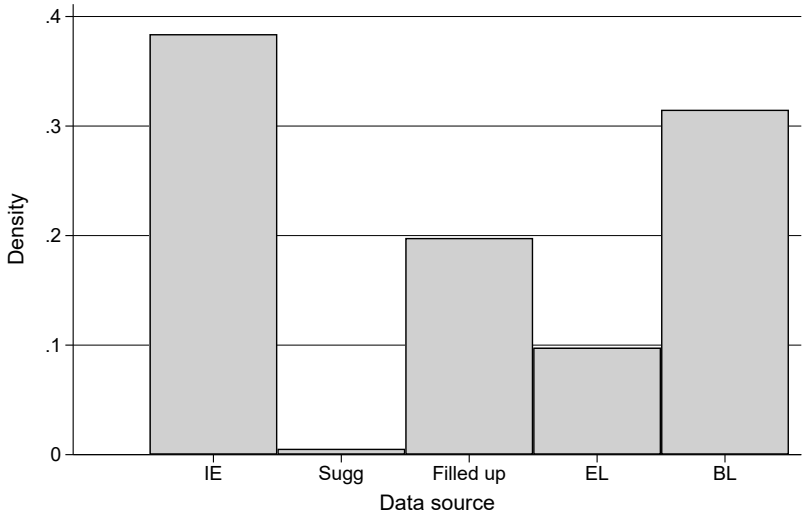


Figure B.1: Data sources for line location of workers

Notes: Graph B.1 shows the distribution of the different data sources to determine the line location of workers. "IE" means line movement data from Industrial Engineers. "Sugg" means line data from the suggestion papers (only if no IE data available). "Filled up" is missing line movement data which is complemented by available data from dates before or after. "EL" indicates that the line code is taken as provided in the endline survey, and "BL" as provided in the baseline survey (if no IE or Suggestion data available).

C Implementation details and documentation

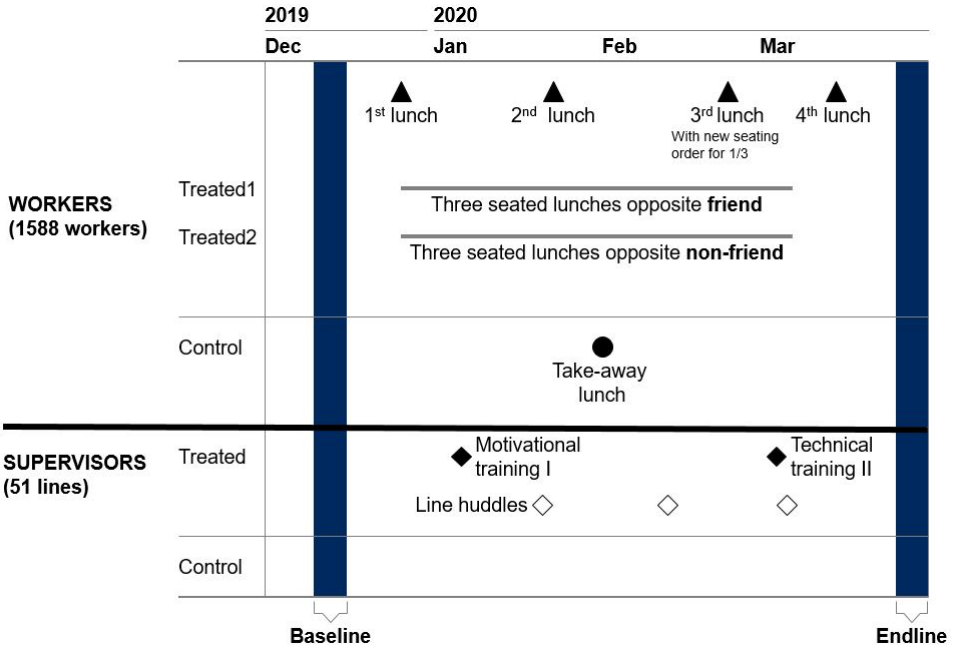


Figure C.2: Timeline

Notes: Figure C.2 shows the timeline of the project between December 2019 and March 2020. Baseline surveys were conducted in mid-December 2019, and endline surveys at the end of March 2020. Every worker assigned to the lunch intervention was supposed to receive lunch once per month over the course of five months. The initial design had to be shortened due to Covid-19, and workers participated in either three or four lunches. The first lunch took place on 26th December 2019. The majority of workers received three lunches between January and March 2020, with only the first of four sub-groups receiving a fourth lunch. I define the time between a lunch of sub-group one and the day before the subsequent lunch of sub-group one as a "round". In each round, each of the four sub-groups received one lunch. The exact lunch dates and rounds can be found in panels a and b of Table C.1. Absent workers received lunch with the next sub-group one week later. The control group was given a take-away lunch in two cohorts on the 21st and the 26th of January 2020 and on the 3rd and 12th of February in case they were absent in January. The motivational training for supervisors took place on the 6th of January 2020 and the technical training on the 4th of March 2020. The first SMS reminder was sent out on the 13th of January (see panel c of Table C.1), with the first huddle scheduled to take place on the 14th of January. Hence, the second managerial intervention started later than the first managerial intervention. In regressions estimating the effect of the second managerial intervention or of both interventions jointly, I drop all observations prior to the 14th of January 2020.

TABLE C.1. INTERVENTION DATES

Panel (a): Dates of peer-to-peer lunch sessions for treatment group (lunch intervention)	
Round 1: First lunch - Subgroup 1	26.12.2019
Round 1: First lunch - Subgroup 2	01.01.2020
Round 1: First lunch - Subgroup 3	08.01.2020
Round 1: First lunch - Subgroup 4	15.01.2020
Round 2: Second lunch - Subgroup 1	23.01.2020
Round 2: Second lunch - Subgroup 2	30.01.2020
Round 2: Second lunch - Subgroup 3	05.02.2020
Round 2: Second lunch - Subgroup 4	12.02.2020
Round 3: Third lunch - Subgroup 1	24.02.2020
Round 3: Third lunch - Subgroup 2	27.02.2020
Round 3: Third lunch - Subgroup 3	03.03.2020
Round 3: Third lunch - Subgroup 4	11.03.2020
Round 4: Fourth lunch - Subgroup 1	15.03.2020
Panel (b): Dates of take-away lunch distribution for control group (lunch intervention)	
Take-away lunch (in round 2) - Subgroup 1	21.01.2020
Take-away lunch (in round 2) - Subgroup 2	26.01.2020
Take-away lunch (in round 2) - Absent workers 1	03.02.2020
Take-away lunch (in round 2) - Absent workers 2	12.02.2020
Panel (c): Dates of training and sms reminders (huddle intervention)	
Motivational training	06.01.2020
Reminder I (2 sms)	13.01.2020
Reminder II (2 sms)	08.02.2020
Technical training	04.03.2020
Reminder III (2 sms)	09.03.2020

First sms: "Dear Supervisors, Thank you for participating in the suggestion box project. Please remember to gather all workers on your line together for 5 minutes and encourage them to submit ideas. Please make sure to meet between XX.XX.2020 and XX.XX.2020 [*two days after reminder I and three days after reminder II and III*]. Also remember to fill out the meeting form that you can get from the floor IE."
Second sms: "Dear Supervisor, You are one of the only supervisors who have not met your workers yet for the suggestion box project. Please be reminded to gather all workers on your line together for 5 minutes and encourage them to submit ideas. Please make sure to meet by tomorrow and remember to fill out the meeting form that you can get from the floor IE. Many thanks."

Figure C.3: SMS texts for supervisors

IDEA BOX (1 idea per paper)

Please write legibly.

Name		**Worker ID**	**Line code**
Department	**Unit**	**Designation**	**Date** ____/____/20____
Problem	**Idea to solve problem**		

Area of improvement – Please tick the box(es) that match your idea

☐ Absenteeism ☐ Defects ☐ Efficiency ☐ Machines ☐ Motions/Waiting time ☐ Production flows/layout

☐ Safety ☐ Shipment ☐ Skills ☐ Wastage ☐ Work environment ☐ Other (please specify): _____

Figure C.4: Suggestion paper

Suggestion Box Line Huddle

Date ……. /……. /2020

Duration of huddle in minutes ……

Your name ……

Your ID ……

Your line and unit ……

No. of line workers ……..

No. of present workers at huddle……..

Approx. no. of ideas from last week ……

Topics discussed

☐ Remind workers of idea box ☐ Motivate workers ☐ Provide help and answer questions ☐ Explain how to fill out idea paper

Which questions were you asked?

Signature Supervisor …….. Signature In-charge ……..

Figure C.5: Template for line supervisors to track conduct of huddles

Suggestion Box – Evaluation Committee
Nomination form

Please select your unit
☐ Unit 1 ☐ Unit 2 ☐ Unit 3
☐ Unit 4 ☐ Unit 5 ☐ Unit 6

Date ……. /……. /2020

What are the best three suggestions?

NO.	IDEA	REASONS FOR SELECTION
#1. Gold		
#2. Silver		
#3. Bronze		

Name, Signature

IE ……, …… Sewing ……, ……

Quality ……, …… Sewing ……, ……

Figure C.6: Template for evaluation committees to track choice of winning ideas

Figure C.7: Sample display of winning suggestions on the floor's notice board

TABLE C.2. TRAINING CURRICULUM

Training type	**Training content**
Motivational training for supervisors (1 hour, 2 groups)	• Details about new suggestion system • Introduction of problem-solving culture • Definition of roles and responsibilities of supervisors – How to maintain good relationships with workers and be a good leader? – How to encourage workers to submit ideas (bi-weekly huddles and on-the-job encouragement)? – How to create an effective and friendly suggestion culture through positive and constructive feedback and recognition?
Technical training for supervisors (1 hour, 2 groups)	• Experience with line huddles • Overview of submitted suggestions • Technical advice – Difference between problem- and solution-oriented submissions – Examples of improvement areas – Typical questions from workers and corresponding answer options
Evaluation committee training (1 hour, 1 group)	• Details about new suggestion system • Success factors of suggestion system • Tasks of evaluation committee – Monitoring of supervisors – Evaluation of suggestions and alignment of three winners with managers – Reward of prizes and display of names – Development of implementation plan

TABLE C.3. OVERVIEW OF HUDDLES

Line code	Floor	Treatment	SMS 1	Date 1st	SMS 2	Date 2nd	Date 3rd	Total	≥ 1	Average across all huddles: Attend. rate	Number workers	Workers per line	Duration in min.	Expect. ideas	Number tasks
1304001	1	1	13.01.2020	03.02.2020	08.02.2020			1	1	100%	25	25	15,0	0	4
1304002	1	1	13.01.2020	14.01.2020	08.02.2020	02.02.2020		2	1	56%	14	25	17,5	1	4
1304003	1	1	13.01.2020	14.01.2020	08.02.2020	03.02.2020		2	1	73%	20	28	15,0	1	4
1304006	1	0				26.02.2020		1	1	100%	29	29	10,0		4
1304008	1	1	13.01.2020	14.01.2020	08.02.2020	03.02.2020		2	1	93%	26	28	15,0	3	4
1304017	2	1	13.01.2020	21.01.2020	08.02.2020	08.02.2020		2	1	64%	35	52	10,0	11	4
1304020	2	1	13.01.2020		08.02.2020	10.02.2020		1	1	43%	25	58		12	4
1304021	2	0				11.02.2020		1	1	51%	20	39	10,0	8	4
1304022	2	1	13.01.2020	21.01.2020	08.02.2020	10.02.2020		2	1	77%	33	43	10,0	10	4
1304024	2	1	13.01.2020	14.01.2020	08.02.2020			1	1	75%	30	40	10,0		4
1304034	3	1	13.01.2020	14.01.2020	08.02.2020	11.02.2020	26.02.2020	3	1	99%	28	28	5,0	2	1
1304035	3	1	13.01.2020	14.01.2020	08.02.2020	11.02.2020	26.02.2020	3	1	100%	23	22	5,0	2	1
1304037	3	1	13.01.2020	14.01.2020	08.02.2020	11.02.2020	27.02.2020	3	1	88%	22	25	11,0	2	1
1304038	3	1	13.01.2020	14.01.2020	08.02.2020	11.02.2020	27.02.2020	3	1	100%	26	26	4,0	1	1
1304049	4	1	13.01.2020		08.02.2020			0	0						
1304051	4	1	13.01.2020	15.01.2020	08.02.2020			1	1	100%	27	27	5,0	80	4
1304053	4	1	13.01.2020	14.01.2020	08.02.2020	27.02.2020		2	1	95%	26	27	5,5	30	2
1304054	4	1	13.01.2020	19.01.2020	08.02.2020	11.02.2020	27.02.2020	3	1	95%	29	31	8,7	37	4
1304065	5	1	13.01.2020	14.01.2020	08.02.2020			1	1	100%	46	46	10,0	6	4
1304067	5	1	13.01.2020	14.01.2020	08.02.2020	12.02.2020		2	1	56%	23	40	10,0	6	2
1304068	5	1	13.01.2020	14.01.2020	08.02.2020	16.02.2020		2	1	79%	31	38	10,0	2	0
1304070	5	1	13.01.2020	14.01.2020	08.02.2020			1	1	100%	47	47	7,0	3	1
1304072	5	0				16.02.2020		1	1	98%	41	42	10,0	5	0
1304093	6	1	13.01.2020	14.01.2020	08.02.2020	12.02.2020		2	1	90%	25	28	12,5	1	4
1304094	6	1	13.01.2020	14.01.2020	08.02.2020	26.02.2020		2	1	100%	24	24	10,0		4
1304096	6	1	13.01.2020	15.01.2020	08.02.2020	24.02.2020		2	1	100%	29	29	10,0	2	1
1304098	6	0				12.02.2020		1	1	100%	27	27	12,0	0	3
1304099	6	1	13.01.2020	14.01.2020	08.02.2020			1	1			23	10,0		4
1304102	6	1	13.01.2020	14.01.2020	08.02.2020			1	1	100%	23	23	1,2		4
Mean										**86%**	**28**	**33**	**9,6**	**10**	**3**
Nb. lines		**29**		23		21	5								
% of all lines				79%		72%	17%								
Nb. treated lines		**25**		23		17	5								
% of treated lines				92%		68%	20%								

Notes: The data in Table C.3 is based on the huddle forms filled out by each supervisor at the end of a line huddle. "Treatment" is an indicator equal to one if the respective line was randomly assigned to treatment. "Total" is the total number of huddles conducted per line during the intervention period, and "≥ 1" is an indicator for at least one huddle. "Attend. rate" is the average attendance rate across all huddles, which is the number of workers attended ("Nb workers") divided by the number of workers per line ("workers per line"). The latter is taken from the baseline data. "Duration in min." shows the average duration of one line huddle per line. "Expect. ideas" is the average number of ideas a supervisor expects to be submitted by workers on his line. An empty cell means that the supervisor did not fill out this information. "Nb tasks" is the average number of tasks a supervisor has undertaken per line huddle. Supervisors could select up to four tasks: Reminding workers of the suggestion box, motivating workers to submit ideas, explaining how to fill out a suggestion paper, and answering questions of workers.

TABLE C.4. OVERVIEW OF WINNING SUGGESTIONS

Line	Prize	Date of reward	Date of submission	Source	Category	Subcategory	Suggestion summary	Suggestion text (shortened)	Type	Total points
1304040	1	20.02.2020	26.12.2019	lunch	physical work env.	workwear or comfort	provide uniform/mask/id	Everyone should follow the rules and regulations in our factory. We should wear personal masks with our ID to be healthy and increase our output.	idea	16
1304023	1	20.02.2020	26.12.2019	lunch				We need to come to the office on time, avoid excessive talking and jealousy, and work together.	general observation	0
1304096	1	20.02.2020	08.01.2020	lunch	facilities/ convenience	transportation	provide transportation	Due to unavailability of life and transport facility, we can't make it on time to work. We should get both the facilities considering our timely arrival at work.	idea	16
1304055	1	20.02.2020	08.01.2020	lunch	health and safety	training	provide training/ training center	Skilled manpower is required to speed up the work, increase effectiveness and shipment on time. Company should provide training for the best outcome.	idea	18
1304068	1	20.02.2020	15.01.2020	lunch	skills	skill assignments	assign right people in right position	Workers should be assigned tp work based on their skills.	idea	21
1304017	1	22.03.2020	30.01.2020	lunch	psychological/ behavioral	workers behaviour	come to office on time/be attentive	We should come to office on time, be attentive, ensure work quality and not be absent without reason. We should also wear ID cards and masks to ensure safety.	idea	22
1304038	1	22.03.2020	05.02.2020	lunch	efficiency	production targets	less wastage of materials	We should reduce wastage while working.	problem/ behaviour	0
1304097	1	22.03.2020	05.02.2020	lunch	machine	functioning	machine malfunctioning	It causes problems if the needle breaks. We should be more attentive and manage time while working.	problem/ behaviour	0
1304071	1	22.03.2020	06.02.2020	box				Everyone should follow company rules, work attentively, reduce higher absenteeism, do the work they are good at and work together to enhance production.	general observation	0
1304021	1	22.03.2020	27.01.2020	box	psychological/ behavioral	workers behaviour	be attentive	We should work attentively to increase efficiency and ensure security. We should also reduce wastage while working.	idea	22
1304001	1	22.03.2020	12.02.2020	lunch				Extra clothes should be provided to reduce defects in finished goods.	general observation	0
1304040	2	20.02.2020	26.12.2019	lunch				I don't have any problem since the mechanics are provided when required.	no problem	0
1304018	2	20.02.2020	01.01.2020	lunch	psychological/ behavioral	workers behaviour	come to office on time	We should come to work on time and work properly to improve quality.	idea	22
1304066	2	20.02.2020	01.01.2020	lunch	absenteeism	attendance bonus	provide attendance bonus	Our attendance bonus should be paid.	idea	18
1304001	2	20.02.2020	25.12.2019	box	skills	operator skills	lack of skilled workers	Production and quality are getting hampered by the lack of skilled workers, hence there is a delay in product delivery.	problem/ behaviour	0
1304055	2	20.02.2020	02.01.2020	box	absenteeism	leave	provide/ grant leave	We should be provided with our leaves and attendance bonus.	idea	23
1304100	2	20.02.2020	15.01.2020	lunch	absenteeism	leave	provide/ grant leave	Our leaves need to be granted to reduce absent rate in factory.	idea	23
1304050	2	22.03.2020	23.01.2020	lunch	manpower	lack of general manpower	provide manpower	Sufficient manpower should be given in each line and they should work attentively and efficiently to fulfill their targets.	idea	18
1304039	2	22.03.2020	23.01.2020	lunch	psychological/ behavioral	workload and pressure	work in all lines/no fixed line	We don't have any fixed line to work at, hence the production rate and quality is low.	problem/ behaviour	0
1304019	2	22.03.2020	23.01.2020	lunch	psychological/ behavioral	workers behaviour	be attentive	We should work attentively to have a quality output. Everyone should work together and the supervisors should motivate the workers.	idea	22
1304070	2	22.03.2020	30.01.2020	lunch	absenteeism	attendance bonus	provide attendance bonus	We should be provided with attendance bonus.	idea	18
1304100	2	22.03.2020	30.01.2020	lunch	absenteeism	leave	provide leave if someone is ill	We should get leaves when we are ill so that we can work in good health and be motivated towards work.	idea	22
1304002	2	22.03.2020	03.02.2020	box	psychological/ behavioral	workers behaviour	be attentive	We need to work attentively to avoid alters and defects. The critical works should be closely monitored.	idea	22
1304036	3	20.02.2020	01.01.2020	lunch				We should all work together to imrove the work environment.	general observation	0
1304068	3	20.02.2020	01.01.2020	lunch	absenteeism	leave	provide/ grant leave	Our leaves should be granted so that we don't remain absent.	idea	23
1304004	3	20.02.2020	25.12.2019	box	absenteeism	attendance bonus	provide attendance bonus	Percentage of absenteeism will decrease and productivity will increase if we get attendance bonus and leaves.	idea	18
1304022	3	20.02.2020	15.01.2020	lunch	facilities/ convenience	place for prayer	no place for prayer	We don't have a perfect place for prayers. Work environment can be improved if supervisors show us sympathy and help us.	problem/ behaviour	0
1304038	3	20.02.2020	15.01.2020	lunch	physical work env.	air quality	bad odor from toilet	Bad odor comes from toilet near the input line.	problem/ behaviour	0
1304099	3	20.02.2020	15.01.2020	lunch	absenteeism	leave	provide/ grant leave	Absenteeism will be reduced if we are granted our leaves timely.	idea	23
1304053	3	22.03.2020	23.01.2020	lunch	machine	availability	availability of machines	We face problems due to machine malfuntioning, line changing and not granting any leaves.	problem/ behaviour	0
1304097	3	22.03.2020	23.01.2020	lunch	absenteeism	attendance bonus	provide attendance bonus	We will be benefited if attendance bonus and sufficient helpers are provided. We can give full effort if we are not scolded.	idea	18
1304069	3	22.03.2020	30.01.2020	lunch	motion waiting time	movement of workers	less movement of workers	Excess movement of workers need to be reduced.	problem/ behaviour	0
1304040	3	22.03.2020	25.01.2020	box	physical work env.	cleanliness	use machine cover while dusting	Machines should be covered while dusting and cleaning should be done after working hours to avoid dirt.	idea	21
1304021	3	22.03.2020	27.01.2020	box	absenteeism	leave	provide/ grant leave	Efficiency will increase if we get our leaves, work attentively and keep safety measures.	idea	23
1304001	3	22.03.2020	12.02.2020	lunch	absenteeism	leave	provide/ grant leave	Rate of absenteeism will decrease if leaves are granted. We should come to the office on time and supervisors should solve our problems.	idea	23

Notes: Table C.4 gives an overview of all winning suggestions which have been selected by the evaluation committees over the course of the interventions. A prize of one is rewarded with 500Tk., a prize of two with 300Tk. and a prize of three with 100Tk. Suggestions can be from the suggestion box (box) or from the active managerial intervention (lunch). Type and total points are assessments by an industry expert, who is independent from the factory.

D Sixth-step randomisation approach into lunch treatment

The randomisation of 1.599 workers into lunch and non-lunch groups was implemented at the pair level (dyads of workers) to have a similar number of friend and non-friend dyads in both groups. Dyads in the control group were never paired in reality. I carried out the randomisation in six steps as described below.

i. Split of sample into workers with and without friends: To begin with, I split the sample into workers with friends on the same production floor (sample one) and without friends (sample two). The first sample (sample one) consists of 1439 workers with friends on the same floor. 714 have one friend and 725 have two friends on the same floor. The second sample (sample two) consists of 160 workers. 51 of them have friends on a different floor (group one), 80 have no friends but are mentioned by someone else (group 2), and 39 have no friends and are also not mentioned by anyone else on the same floor. The numbers are also displayed in Table D.5. The reason for these categorizations is that workers with friends might have different characteristics or behaviours than workers without. By excluding the latter, I can test whether results still hold, in which case they would not behave differently.

TABLE D.5. OVERVIEW OF TWO RANDOMISATION SAMPLES

Total number of workers	**1599**
SAMPLE 1	**1439**
- with one friend on the same floor	714
- with two friends on the same floor	725
SAMPLE 2	**160**
- group 1: with friends on a different floor	41
- group 2: without friends (mentioned by someone else on the same floor)	80
- group 3: without friends (not mentioned by someone else on the same floor)	39

ii. Generation of dyads: As a next step, I generate pairs of workers. This step is first done for workers with friends (sample one), and then for workers without friends (sample two). There are two instances, in which workers are moved from sample two into sample one, or the other way round. On the one hand, some workers in sample two, which do not mention any friends, but are mentioned by someone else as a friend, could be 'drawn' into sample one as I assume that friendships are symmetric. On the other hand, some workers in sample one could become part of sample two if they are not able to be paired, which is depending on the density of the network.

Dyads of friends: In order to generate as many dyads of friends as possible in a scarce network, I pair workers in a particular order depending on the number of friends and the number of times a worker is mentioned by someone else as a friend. I first pair workers with only one friend and then workers with two friends. Within both groups, I first pair workers who have been mentioned least by someone else as a friend and lastly pair workers who have been mentioned most by someone else, i.e. I start with the most unpopular workers (zero times mentioned by someone else) and end with the most popular workers (eleven times mentioned by someone else). The distribution of the number of times being mentioned as a friend is shown in Figure D.8. As the number of possible matches for unpopular workers is lower than for popular workers, this approach allows us

to come up with the highest number of possible matches in a scarce network. Workers who cannot be paired are added to sample two (group 4: 'leftovers').

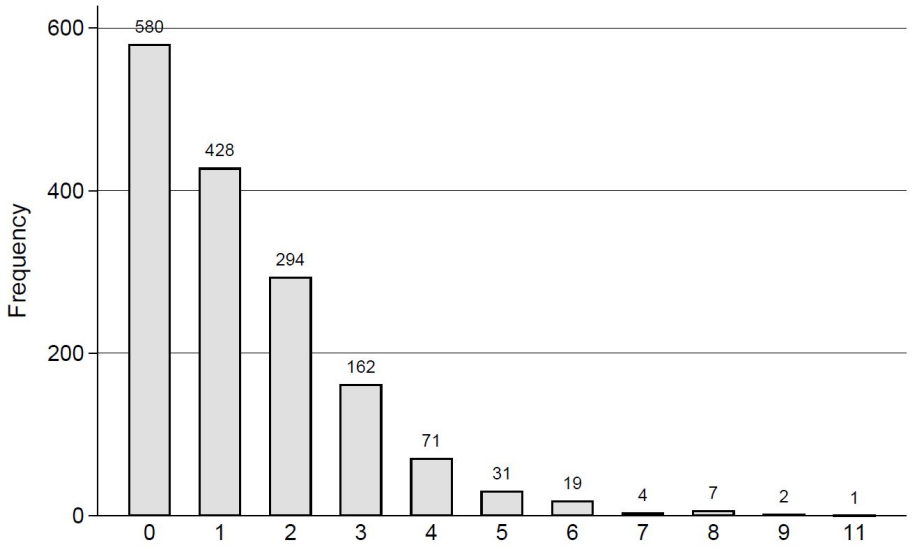


Figure D.8: Frequency of workers by number of times being mentioned as friend

Dyads of non-friends: I generate as many dyads of non-friends as possible for workers in sample two, i.e. four sub-groups of workers: Workers with friends on different floors (group 1), workers without friends but mentioned by someone else on the same floor (group 2), workers without friends but not mentioned by anyone on the same floor (group 3), and workers who could not be paired in the previous step (group 4). The matching process is randomly performed by floor and by sub-group. Very few workers are not able to be paired at all.

iii. Randomisation into lunch and non-lunch group: I now randomise the dyads of friends/non-friends into treatment and control group by sample such that sample one and sample two consists of a balanced number of workers in both groups. I hereby take into account whether a dyad consists of at least one worker who had been asked an extended survey question or not. This ensures to have a balanced number of extended survey workers in both groups for each line. Hence, I first generate dyads with at least one extended survey worker stratified by sample and by line. All uneven remaining dyads are randomly assigned by sample only. I then generate dyads with no extended survey workers stratified by sample and by floor. All uneven remaining dyads are completely randomly assigned.

iv. Re-balancing: I re-balance treatment assignments to achieve balance on the maximum experience index per dyad across treatment groups and on the percentage of treatment dyads across lines. To do so, I repeat above mentioned steps ii. and iii. 100 times and randomly select one assignment which fulfills the following two criteria:

- The p-value of the differences in maximum experience indices per dyad should be larger than 0.2. I select the maximum experience index per dyad as I expect the performance of a pair to be determined by the highest experience index, and not by the average.

- The top decile of the difference in the number of treatment and control dyads per line (as a percentage of all dyads per line) of sample 1 should be < 0.45 and of sample 2 < 0.75. Effects might be different if a supervisor is exposed to fewer or more lunch workers. It would be interesting to exploit that difference, but I do not have enough power to do that.

v. Re-pairing of friend dyads: As a penultimate step, I randomly select 50% of the dyads (out of the friend dyads from sample one) to be not seated with a friend. These dyads are taken apart and workers are re-paired with a non-friend on the same floor. This step is done to be able to compare friend and non-friend dyads within sample one, i.e. within workers who have friends. The final number of workers per treatment assignment is shown in Table D.6.

TABLE D.6. TREATMENT ASSIGNMENTS

SAMPLE 1	**1244**
Workers in treatment group (friends and non-friend dyads)	**624 (312 pairs)**
- hereof workers in friend dyads	310 (155 pairs)
- hereof workers in non-friend dyads	314 (157 pairs)
Workers in control group	**624 (312 pairs)**
SAMPLE 2	**340**
Workers in treatment group (non-friend dyads only)	**170 (85 pairs)**
– group 1: friends on different floor	16
– group 2: without friends (mentioned by someone else on the same floor)	8
– group 3: without friends (not mentioned by someone else on the same floor)	18
– group 4: leftovers from sample one	128
Workers in control group	**170 (85 pairs)**
– group 1: friends on different floor	20
– group 2: without friends (mentioned by someone else on the same floor)	8
– group 3: without friends (not mentioned by someone else on the same floor)	18
– group 4: leftovers from sample one	126
Not in any sample / unpaired	**11**

vi. Change of seating order: After lunch rounds one and two, I randomly select 1/3 of all pairs assigned to lunch – 1/3 of the friend and non-friend pairs in sample one and 1/3 of the non-friend pairs in sample two – to be seated opposite a new non-friend, as shown in Table D.8. I first dropped all workers who have left the factory, and indicate whether a worker's partner has left (see Table D.7).

TABLE D.7. NUMBER OF EXITS BEFORE CHANGE OF SEATING ORDER

Workers in treatment group	795
- hereof left	104
Remaining	**691**
Workers in control group	794
- hereof left	82
Remaining	**712**

I then balanced on an indicator equal to one if a worker's seating partner had left the factory by the time the change of the seating partners was arranged, i.e. I randomly chose 1/3 of workers whose partner had left, and 1/3 of workers whose partner had not left. Next, I re-paired all those workers within sample one and within sample two with a non-friend by floor. Finally, I randomly chose a seating order for the dyads. Workers whose partner had left (but have not been re-paired) were randomly seated at the edges of the tables, because I could not leave empty spaces at a table. The number of workers who had new seating partners per sample are shown in Table D.8.

TABLE D.8. CHANGE OF SEATING ORDER

SAMPLE 1	549
Seating change	182
- *hereof friend dyads*	90
- *hereof non-friend dyads*	92
Unchanged	367
- *hereof friend dyads*	181
- *hereof non-friend dyads*	186
SAMPLE 2	141
Seating change	49
Unchanged	92

E Data sources

The analysis draws upon four main sources of data: Intervention output, administrative data, survey data and focus groups.

Intervention output: I obtained submitted suggestion papers filled out by workers, huddle forms filled out by supervisors and nomination forms filled out by the evaluation committees. Suggestion papers were collected by the research team every week at the end of each lunch session, and also from the suggestion box located on the production floor. Huddle forms, which give information about the date, the participants and the content of line huddles, had to be filled out by supervisors and handed over to the IE (Industrial Engineering) officer for signed approval. The IE department then delivered the forms to the research team. Nomination forms with information about three nominated winners per floor had to be filled out by the floor-wise evaluation committees and handed over to the management, which then delivered the forms to the research team.

Administrative data: The analysis uses various factory datasets maintained by different departments. The production department provided daily production data including efficiency and quality information. The HR department provided monthly salary data, daily job card data and skills data. The salary data contains information about the workforce, such as the day of joining and the day of exiting the factory. The daily time stamps of each worker with information about the time of entering and of leaving the factory each day is captured in the job card data. The skills matrix data contains information about each process a worker can perform and the time it takes to perform the respective process.[49] The IE department provided line movement data with the daily working

[49]The IE team records this data using a stopwatch when a worker joins the factory.

location of each worker. Due to incompleteness of the data[50], I draw upon intervention and survey data to complement the line location of workers (see detailed description in Appendix Section B).

Survey data: The analysis also uses survey data of line operators and line supervisors on treatment and control lines collected at two points in time: Shortly before the experiment in early December 2020 (baseline) with the help of eleven enumerators, and shortly after the experiment in mid-March 2020 (endline) with the help of twelve enumerators. The baseline survey consists of one in-person part inside the factory and another phone part. The in-person survey was used to ask about demographics and characteristics, perceptions about the suggestion box and the suggestion culture. The phone survey was used to ask more delicate questions about job satisfaction, feelings and fears. The endline survey was conducted on the phone only, as enumerators were not able to visit the factory due to Covid-19. Both the baseline and the endline survey for line operators consist of a short version for all workers, and a long version for randomly selected workers. For supervisors, there was only one version.

Focus groups: I organised three focus groups with workers, lower/middle managers and higher-level managers each in November 2019 at the end of the pilot experiment. The focus group with workers was comprised of eight workers from the treatment and eight workers from the control group. An IE officer, quality inspector, line supervisor and line chief participated in the focus group with lower/middle managers. The focus group with higher-level managers included four managers, i.e. the General Managers from the Production and the IE department, and the Deputy General Managers from the Quality and the HR department. The purpose of these focus groups was to obtain qualitative, descriptive information on the constraints for workers to submit suggestions and how they felt about the lunch sessions and involvements of supervisors. Although this is a small sample, the focus groups provide richness and a deeper understanding into the mechanisms through which the treatment is expected to have an impact.

F Local average treatment effects

I estimate local average treatment effects (LATE) using the randomised assignment to treatment to instrument for actually receiving the treatment. I use random assignment to the lunch treatment as an instrument for actual participation at at least one lunch, and the random assignment to the huddle treatment as an instrument for actual participation at at least one line huddle. The participation at line huddles is calculated by drawing upon two data sources. First, the exact dates of conducting a huddle per line are identified using information from line huddle templates which supervisors had to submit after each huddle. Second, the exact line position of each worker per day is identified using information from the line movement data. Hence, I infer whether a worker was working on a line on which a line huddle was conducted on a particular day. I then generate a cumulative actual line huddle indicator which is equal to one if a worker has participated at at least one line huddle in the past. Depending on the time structure of the data, I collapse the cumulative actual line huddle indicator by the respective time periods, e.g. by week, month or round. For production data at the line-level, I use the intensity of treatment per line as the actual

[50] 60% of the days during the intervention period are missing, and 159 out of 1,599 workers are not included in the data at all.

treatment. It is calculated by dividing the total number of treated workers by the total number of all workers per line and day.

G Components of main outcomes

TABLE G.9. PRIMARY OUTCOME VARIABLES

PRIMARY OUTCOMES RELATED TO FIRM PERFORMANCE		
WORKFORCE INDEX		
absenteeism	Ratio: number of workers absent / number of total workers (present + absent)	daily clock data
punctuality	Minutes worker arrives before production start date	daily clock data
working hours	Working hours in minutes	daily clock data
Other workforce-related variables (not included in index)		
absenteeism on treatment day	Ratio: number of workers absent on treatment day / number of total workers on treatment day	daily clock data
exits	Ratio: number of workers left / number of total workers	monthly salary data
PRODUCTION INDEX		
efficiency	Ratio: output minutes (smv*daily output) / number of workers	weekly production data
alteration rate	Ratio: number of alterations / number of pieces checked	weekly production data
reject rate	Ratio: number of rejects / number of pieces checked	weekly production data
PRIMARY OUTCOMES RELATED TO WORKER WELL-BEING		
JOB SATISFACTION INDEX (WORKERS)		
Job satisfaction subindex		

Continued on next page

TABLE G.9 – continued from previous page

own job satisfaction	Own job satisfaction (scale 1-4)	survey data
others' job satisfaction	Co-workers' job satisfaction (scale 1-4)	survey data
career opportunities	Satisfaction with career opportunities at work (scale 1-4)	survey data
job training & support	Satisfaction with level of job training and support (scale 1-4)	survey data
Turnover intentions subindex		
intent to quit	Number of times thinking about quitting your job in the last three weeks	survey data
intent to stay	Number of days you would like to continue working in this factory	survey data
Job attitudes subindex		
own attitude about job	Own attitude to do the best work you can, even if it sometimes does interfere with the rest of my life (scale 1-3)	survey data
others' attiude about job	Other operators' attitude to do the best work they can, even if it sometimes does interfere with the rest of their life (scale 1-3)	survey data
sense of punctuality	Minutes worker should arrive before work start	survey data

Job satisfaction of supervisors (not included in index)		
Job satisfaction subindex		
own job satisfaction	Satisfaction with own job (scale 1-4)	survey data
operators' job satisfaction	Satisfaction of operators with their job (scale 1-4)	survey data
happiness at work	Happiness while at work (scale 1-10, with ten being very happy, five being sometimes happy and sometimes unhappy and 1 being very unhappy)	survey data
Turnover intentions subindex		
intent to quit	Number of times thinking about quitting your job in the last three weeks	survey data
intent to stay	Number of days you would to continue working in this factory	survey data

Continued on next page

TABLE G.9 – continued from previous page

MENTAL HEALTH INDEX		
Feelings at work subindex		
feeling angry	Number of times you feel angry, hurt, stressed, or frustrated because of a situation that arose at work	survey data
person involved	Number of persons involved in causing the mood	survey data
feeling happy	Number of times you feel happy, cared-for, or good because of a situation that arose at work	survey data
person involved	Number of persons involved in causing the mood	survey data
General anxiety disorder subindex		
nervous	Feeling nervous, anxious, or on edge	survey data
sleepless	Not being able to sleep or control worrying	survey data
worrying	Worrying too much about different things	survey data
trouble relaxing	Trouble relaxing	survey data
restless	Being so restless that it is hard to sit still	survey data
annoyed	Becoming easily annoyed or irritable	survey data
afraid	Feeling afraid, as if something awful might happen	survey data
Workplace harassment subindex		
frequency of abusive language	Number of days other operators on your line needed to put up with shouting or abusive language at work over the last two weeks (scale 1-4)	survey data
frequency of uncomfortable situations	Number of days other operators on your line had to experience situations at work that made them feel uncomfortable, for example persistent staring, flirting, grabbing their hand or other parts of the body, etc. (scale 1-4)	survey data
touched by primary supervisor	Dummy: Primary supervisor every physically touched you in aggressive way	survey data
touched by any supervisor or manager	Dummy: Any supervisor or manager ever talked to you or touched in a sexual way	survey data

PRIMARY OUTCOMES RELATED TO INNOVATION		
quantity	Indicator equal to one for any submitted suggestion and equal to zero for no submitted suggestion per round	intervention output

Continued on next page

TABLE G.9 – continued from previous page

type	Type of suggestion from one to six: 1 incoherent, 2 general observation/no problem, 3 poorly communicated problem, 4 well communicated problem, 5 poorly communicated idea, 6 well communicated idea	intervention output
quality of ideas	Quality index for ideas only (type 5 and 6) consisting of six components on a scale from one to five each (apart from component four which only has a scale of one to two): 1. **Implementation level:** At which level can the idea be best implemented? (one workplace, specific workplaces, line level, floor level, or factory level) 2. **Impact on worker well-being:** How large would be the expected benefit for affected workers? (very low to very high) 3. **Impact on efficiency:** How large would be the expected effect on productivity? (very low to very high) 4. **Level of effects:** Are productivity effects direct (e.g. repaired machine) or indirect through improved motivation of workers (e.g. prayer place)? 5. **Implementation difficulty:** How fast and riskless is the suggestion to implement? (very long to very fast) 6. **Financial resources:** How much money is needed for implementation? (very substantial over BDT 50k to none)	intervention output

Continued on next page

TABLE G.9 – continued from previous page

quality	Quality index of suggestions ranging from zero to ten, which is an additive combination of the type and - for type five and six (ideas) - the quality of ideas. It is constructed as follows: Types 1 to 4 are not further assessed, and receive one to four points, respectively. Types 5 and 6, i.e. ideas, are further evaluated and assigned between one and five additional points. Thus, type 5 can have between five and nine points, and type 6 between six and ten points in total. The additional points for type 5 and 6 are based on the sum of the five criteria described above (see quality of ideas), which are evaluated on a scale from one to five (apart from one criterion on a scale from one to two). The sum is clustered into five ranges, which make up the additional points. The five ranges are between 1-19, 20, 21-22, 23, and 24-27 points such that each range has a similar number of suggestions. Non-submissions are included with a value of zero	intervention output
quality (if submitted)	Same as quality index above, but non-submissions are excluded, i.e. ranging from one to ten	intervention output
worker components	Same as quality index above, but quality index of ideas only measures impact on worker outcomes (components three and four)	intervention output
firm components	Same as quality index above, but quality index of ideas only measures impact on firm outcomes (components five to eight)	intervention output
content	Indicator equal to one if submitted suggestion is classified as respective category; twelve different indicators for the following categories: Supervisor skills, leave, attendance bonus, work environment (behavioral), work environment (facilities), manpower, machine, salary, motion/waiting time, efficiency, health and safety, and production layout	intervention output

Continued on next page

TABLE G.9 – continued from previous page

variety	Number of submitted suggestions per round which have not been mentioned by worker before	intervention output

TABLE G.10. VARIABLES TO MEASURE HETEROGENEOUS EFFECTS

worker's experience index	Anderson index including number of factories worked at, grade, tenure in factory in years, tenure in industry in years, years of education	survey data
processes	Number of sewing processes a worker can perform	skills data
worker's gender	Indicator equal to one for female worker	survey data
new seating order	Indicator equal to one for workers who were seated next to a new partner in lunch round three	randomisation data
friend as seating partner	Indicator equal to one for workers who were seated opposite a friend	randomisation data
friendship intensity	Number of topics worker discusses with friend and number of locations	survey data
number of friends	Number of friends as seating partners (left, right, opposite, diagonally left, diagonally right)	intervention output
cooperativeness	Similarity measure of suggestions from seating partners	intervention output
supervisor's experience index	Anderson index including number of factories worked at, tenure in factory in years, tenure in industry in years, years of education	survey data
supervisor's gender	Indicator equal to one for female worker	survey data
perceived efficacy supervisors	Beliefs about winning and implementation, and assessment of workers	survey data

Content area	Category	Sub-category	Suggestion paper	Idea	Problem
Work environment	Work environment - behavioral/ psychological	workers behaviour	unnecessary chaos/chatting on floor		✓
			avoid harassment	✓	
			avoid unnecessary talking	✓	
			be attentive	✓	
			come to office on time	✓	
			follow rules and regulations	✓	
		workload and pressure	partial behaviour in workplace		✓
			excessive workload		✓
			work in all lines/no fixed line		✓
			avoid partiality	✓	
			assign working on fixed line	✓	
			leave on time	✓	
			reduce work pressure	✓	
	Work environment - facilities/ convenience	lift	no availability of lift		✓
			technical difficulties of lift		✓
		toilet	not allowed to go to toilet		✓
		place for prayer	not allowed to pray		✓
			no place for prayer		✓
			provide prayer place for girls	✓	
			provide prayer place/mosque	✓	
		punch machine	no punch machine		✓
			provide punch machine	✓	
		transportation	no transportation		✓
			provide transportation	✓	
		road condition	clear road from vehicles to cross safely	✓	
			provide foot overbridge	✓	
		water facility	need heater for warm water	✓	
			need water during lunch	✓	
		lunch	provide lunch from factory	✓	
			allow to bring food on floor	✓	
		yearbook	publish annual/yearbook	✓	
	Work environment - workplace	air quality	bad odor from toilet		✓
			affected by polluting air		✓
		light	not enough light		✓
		temperature	less/no fans		✓
			cold temperature		✓
			wind from outside		✓
			hot temperature		✓
			provide fans	✓	
		cleanliness	dirty working place/toilet		✓
			assure cleanliness of workplace	✓	
			avoid spitting/littering	✓	
			provide soap	✓	
			use machine cover while dusting	✓	
		damages	broken chair/table		✓
			fix broken chair	✓	
			repair damaged floor tiles	✓	
		noise	place soundbox in middle	✓	
			play national anthem in the morning	✓	
			no music after lunch		✓
		workwear or comfort	no shoes allowed		✓
			provide chair/table	✓	
			provide cushion	✓	
			provide shoe boxes	✓	
			provide uniform/mask/id	✓	
		box	provide boxes	✓	
	Salary	base salary	insufficient salary		✓
			need salary on time	✓	
			provide more salary	✓	
			provide salary in hand	✓	
		bonus salary	not enough night bill/overtime		✓
			provide benefits/increment/production bonus	✓	
		fairness	partiality in benefits/increment		✓
			salary discrimination		✓
			need promotion for helpers	✓	
			provide fair salary based on hierarchial level	✓	
			reduce discrimination	✓	
	Health and safety	health tools/equipment	not wearing mask		✓
		pregnancy	many movements for pregnant woman		✓
		safety measures	no CCTV		✓
			provide cctv camera	✓	
			form a safety committee	✓	
		medical care	longer availability of doctors from 8am till night	✓	
			provide eye test facility	✓	
		training	provide training/training center	✓	

Continued on next page

Figure G.9: Categorizations of suggestions

Figure G.9 – continued from previous page

Content area	Category	Sub-category	Suggestion paper	Idea	Problem
Supervisor skills	Supervisor skills	behaviour and tasks	mismanagement in work allocation		✓
			assign right people in right position	✓	
			misbehaved/scolded by supervisors		✓
			inconsiderate/less supportive supervisors		✓
			no guidance in newly assigned work		✓
			work done forcefully		✓
			appoint better supervisor	✓	
			avoid scolding/misbehaviour	✓	
			need more support from supervisors	✓	
			need raw material on time	✓	
			provide educated supervisor	✓	
			provide good behaviour	✓	
		supervisor gender	female supervisors wanted		✓
			female supervisors not wanted		✓
			need male supervisor	✓	
Leave/attendance bonus	Leave	leave	too high absenteeism		✓
			understand reasons for leave/leave not granted		✓
			walk around a lot to request leave		✓
			provide/grant leave	✓	
			provide leave if someone is ill	✓	
	Attendance bonus	attendance bonus	no attendance bonus		✓
			provide attendance bonus	✓	
Production	Salary	base salary	insufficient salary		✓
			need salary on time	✓	
			provide more salary	✓	
			provide salary in hand	✓	
		bonus salary	not enough night bill/overtime		✓
			provide benefits/increment/production bonus	✓	
		fairness	partiality in benefits/increment		✓
			salary discrimination		✓
			need promotion for helpers	✓	
			provide fair salary based on hierarchial level	✓	
			reduce discrimination	✓	
	Motion/ waiting time	non-productive time	false fire alarm		✓
			waiting for partner to come		✓
		movement of workers	less movement across lines and processes		✓
		location of canteen	canteen too far away		✓
			need canteen at nearest location	✓	
	Efficiency	labelling	bundle cut should reach output table		✓
			maintain serial after bundle cut	✓	
			provide correct body stickers	✓	
		production targets	lack of extra bobbin		✓
			less wastage of materials		✓
			need quicker process in production	✓	
			measure production rate	✓	
			need more target	✓	
		quality check	change cutter	✓	
			need to check quality	✓	
	Health and safety	health tools/equipment	not wearing mask		✓
		pregnancy	many movements for pregnant woman		✓
		safety measures	no CCTV		✓
			provide cctv camera	✓	
			form a safety committee	✓	
		medical care	longer availability of doctors from 8am till night	✓	
			provide eye test facility	✓	
		training	provide training/training center	✓	
	Production layout	error	change layout		✓
			shipment problem		✓

Notes: Figure G.9 shows the categorizations of all submitted suggestions into content area, category, subcategory and suggestion paper by an industry expert. A tick in the idea column indicates that respective suggestions were assessed as the type "idea". A tick in the problem column indicates that respective suggestions were assessed as the type "problem".

	Survey components		
	✓	included	
	×	excluded	
Survey questions	**Operators**		**Supervisors**
Futility Index	✓		✓
Self-assessment of own submitted suggestions	✓		×
Likelihood of winning	✓		×
Likelihood of implementation of winners' idea	✓		✓
Likelihood of implementation of own idea	✓		×
Feeling incentivized to submit idea	✓	Endline, all	×
Likelihood of prize on own line	×		✓
Satisfaction with award of prizes	×		✓ Endline
Workers as biggest beneficiary of suggestion box	×		✓ Endline
Motivation of typical worker in factory vs line	×		✓
Concentration of typical worker in factory vs line	×		✓
Time Index	✓		✓
Time as a constraint	✓		✓
Frequency thinking about problems	✓		✓
Frequency thinking about solutions	✓		✓
Thinking about ideas over lunch	✓		×
Frequency of discussing ideas with people outside of the factory	✓		×
Knowledge Index	✓		✓
Knowledge as a constraint	✓		✓
Easiness to think about problems	✓		×
Easiness to think about solutions	✓		×
Ability to develop suggestions of typical worker in factory vs line	×		✓
Security Index	✓		✓
Subindex: Voice climate workers	✓		✓
Sharing of personal issues with superior	✓		×
Operators sharing issues with superior	✓		✓
Importance of sharing personal issues	✓		✓
Ability to be yourself at work	✓		×
Ability to openly state opinions	✓		✓
Fear of doing something wrong	✓		✓
Fear of admitting an error	✓		✓
Misreporting to supervisor	✓		✓
Degree of trust to workers (telling a secret)	✓		✓
Degree of trust to workers (lending money)	×		✓
Feeling valued if suggestion made	✓		×
Feeling comfortable to submit idea	✓	Endline, all	×
Conflict between workers and managers	✓		✓
Operator suggestion culture	×		✓
Feeling responsible for changes on the production line	✓		×
Influence over workers	×		✓
Importance of group goal setting compared to supervisor goal setting	×		✓
Subindex: Leaderhip efficacy	✓		×
Caring culture (index)	✓		×
Feeling listened to	✓		×
Feeling treated with respect	✓		×
Feeling appreciated & recognized	✓		×
Support from others	✓		×
Effort of supervisor to get to know worker	✓		×
Efforts to get to know each other	✓		×
Supervisor abilities (index)	✓		✓
Ability of typical supervisor to encourage operators	✓		✓
Ability of own supervisor/yourself to encourage operators	✓		✓
Difference in ability to encourage operators between typical and own	✓		✓
Overall ability of typical supervisor	✓		✓
Overall ability of own/yourself supervisor	✓		✓
Difference in overall ability between typical and own supervisor	✓		✓
Decision-making authority as a constraint	✓		✓

Figure G.10: Secondary outcome variables

	Survey components			
	✓	included		
	×	excluded		
Survey questions	**Operators**		**Supervisors**	
Communication index	✓		✓	
Subindex: General communication	✓		✓	
Number of topics discussed with supervisor	✓		×	
Something discussed	✓		×	
Work-related grievances/problems discussed	✓		✓	
Improvement of skills discussed	✓		✓	
Production capabilities discussed	✓		✓	
Work environment discussed	✓		✓	
Frequency of discussions / number of workers talked to more than three times over last two weeks	✓		✓	
Forms of discussions (individual meetings, group meetings, spontaneously on the line and/or during breaks)	✓		✓	
Worker involved in initiating the discussion	✓		✓	
Subindex: Communication with managers	×		✓	
Frequency of talking about work-related grievances/problems with manager	×		✓	
Frequency of talking about skills with manager	×		✓	
Frequency of talking about production capabilities with manager	×		✓	
Frequency of talking about work environment with manager	×		✓	
Forms of discussions (individual meetings, group meetings, spontaneously on the line and/or during breaks)	×		✓	
Supervisor involved in initiating the discussion	×		✓	
Encouragement to think about suggestions from managers	×		✓	
Level of information about suggestion box from managers	×		✓	Endline
Subindex: Raising suggestions to supervisors/managers	✓		×	
Problem raised to supervisor	✓		×	
Suggestion made to supervisor	✓		×	
Problem raised to manager	✓		×	
Suggestion made to manager	✓		×	
Subindex: Suggestion-box-specific communication	✓		✓	
Frequency of talking about suggestion box with supervisor	✓	Endline, all	×	
Frequency of line group meetings about suggestion box	✓	Endline, all	×	
Number of workers talked about box over last two weeks	×		✓	Endline
Number of workers talked about box more than three times over last two weeks	×		✓	Endline
Forms of discussions (individual meetings, group meetings, spontaneously on the line and/or during breaks)	×		✓	Endline
Worker involved in initiating the discussion	×		✓	Endline
Content of discussion (organizational issues, content of suggestions, other)	×		✓	Endline
Assessment of group meetings	×		✓	Endline
No concerns such as workers complaining about supervisors	×		✓	Endline
Submission of suggestions which workers would otherwise not tell anyone	×		✓	Endline
Subindex: Submission of suggestions	✓		✓	
Ever submitted an idea into the suggestion box	✓	Endline, all	×	
Ever told an idea to someone	✓	Endline, all	×	
Person to whom idea was told is an operator	✓	Endline, all	×	
Estimated number of suggestions	×		✓	Endline
Frequency of knowing content of suggestions	×		✓	Endline
Knowledge about submitted topics on your line	×		✓	Endline
Own ideas communicated downwards	×		✓	Endline
Own ideas communicated upwards	×		✓	Endline

Figure G.11: Secondary outcome variables

H Supplementary figures

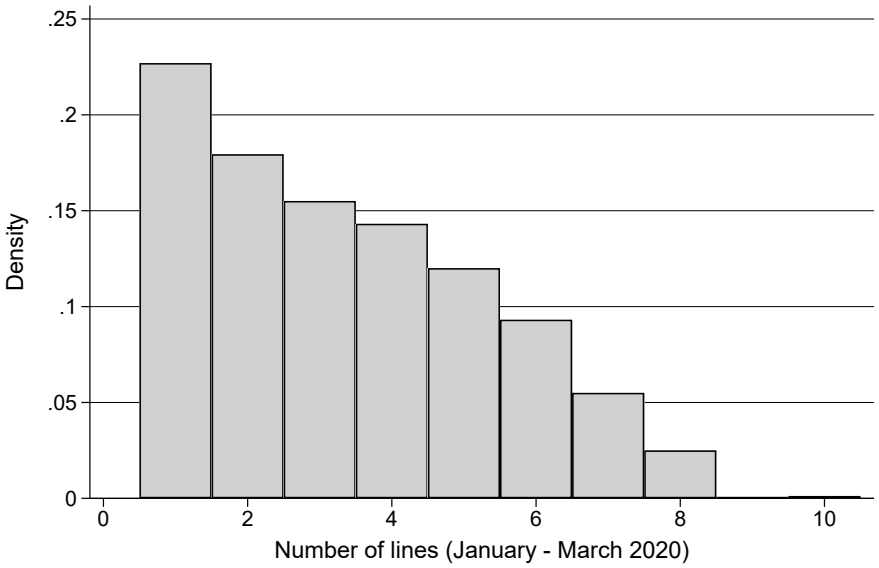


Figure H.12: Distribution of total number of lines per worker

Notes: Graph H.12 shows the distribution of the average number of lines each worker is assigned to over a period of three months, i.e. 91 days, between January 2020 and March 2020.

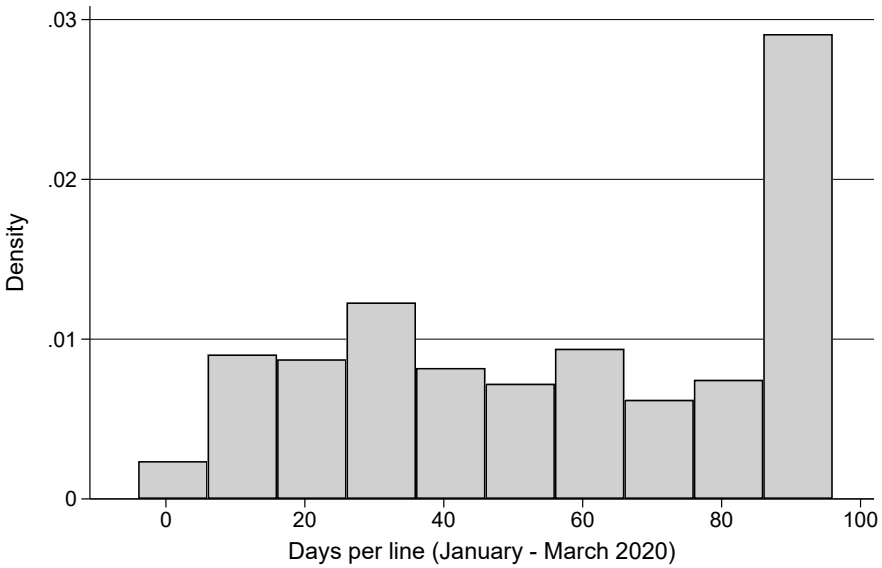


Figure H.13: Distribution of maximum number of days per line per worker

Notes: Graph H.13 shows the distribution of the average maximum number of days each worker is assigned to one line over a period of three months, i.e. 91 days, between January 2020 and March 2020.

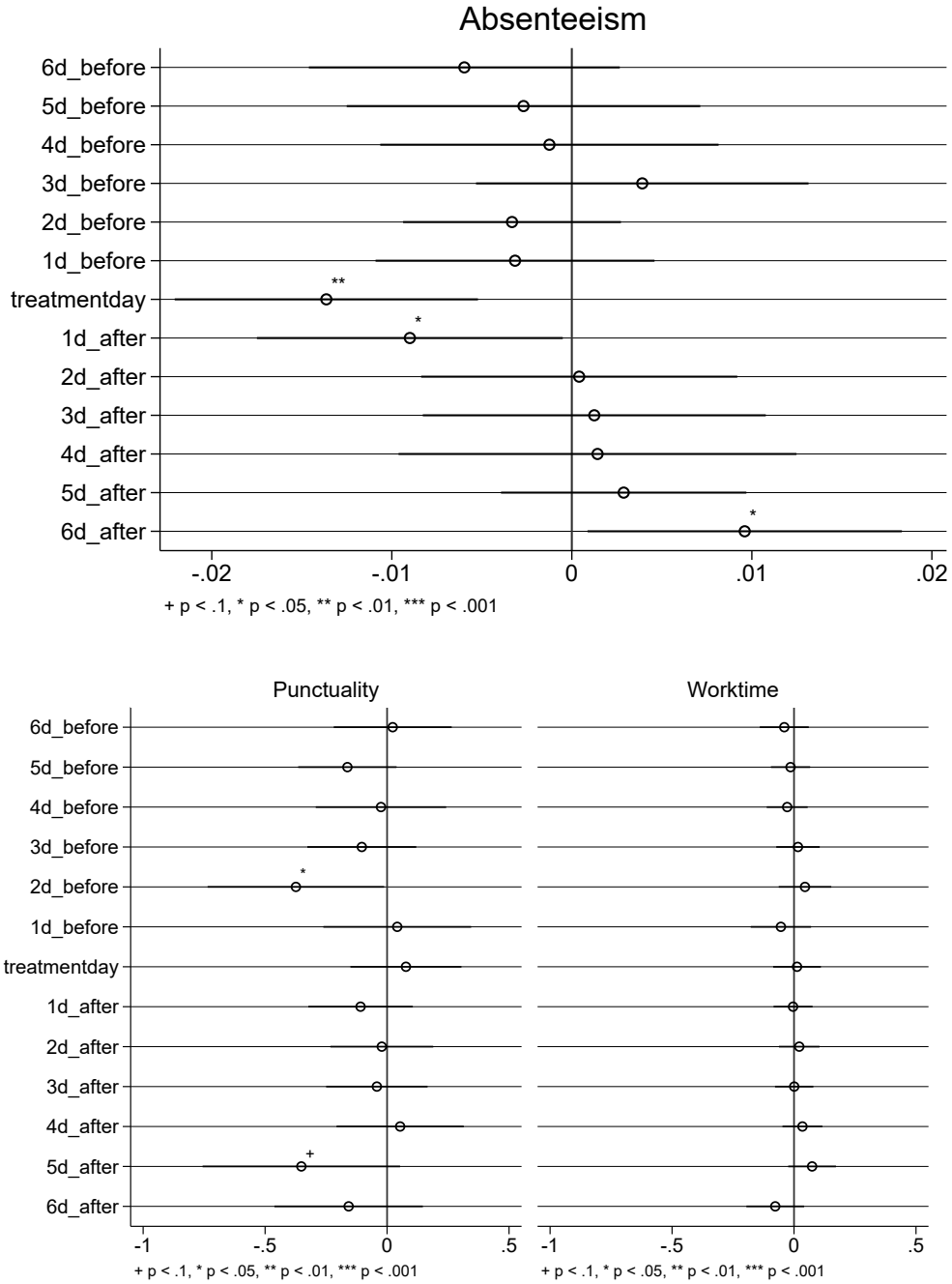


Figure H.14: Daily effects on workforce index components

Notes: Table H.14 shows the plot of coefficients of regressions of the lunch treatment days including six leads and lags on primary outcome variables at the daily level. "treatmentday" is an indicator equal to one on assigned treatment days. "Xd_before" is an indicator equal to one for the Xth day before assigned treatment days, and "Xd_after" for the Xth day after assigned treatment days. Absenteeism is an indicator equal to one if the worker was absent on the respective day. Punctuality measures the time a worker arrives before the assigned shift in minutes. Worktime measures the daily working time in hours.

Figure H.15: Effects on primary outcome variables

Notes: Figure H.15 shows treatment effects on primary outcome variables related to firm performance, represented by the workforce index and the production index, and worker well-being, represented by the job satisfaction index and the mental health index. Indices are based on Anderson (2008). Index components are shown in Table G. Corresponding regression results are shown in Table 1. 95% confidence intervals are displayed.

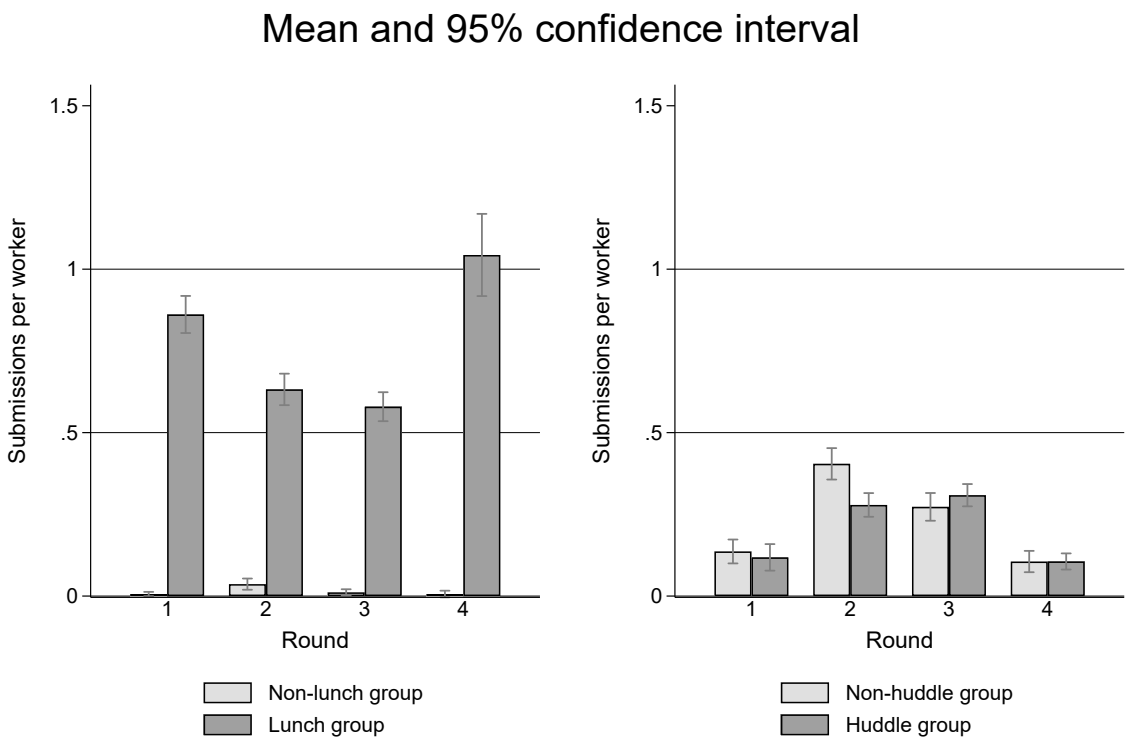


Figure H.16: Average number of suggestions per round

Notes: The left graph of Figure H.16 shows the mean and 95% confidence interval of the number of submitted suggestions by non-lunch and lunch workers per round. The fourth round restricts the data of lunch workers to those who were assigned to group one, in other words who received lunch in round four. The right graph shows the mean and 95% confidence interval by non-huddle and huddle workers. The data excludes all suggestions submitted prior to the date when the first huddle took place, i.e. 14th of January 2020.

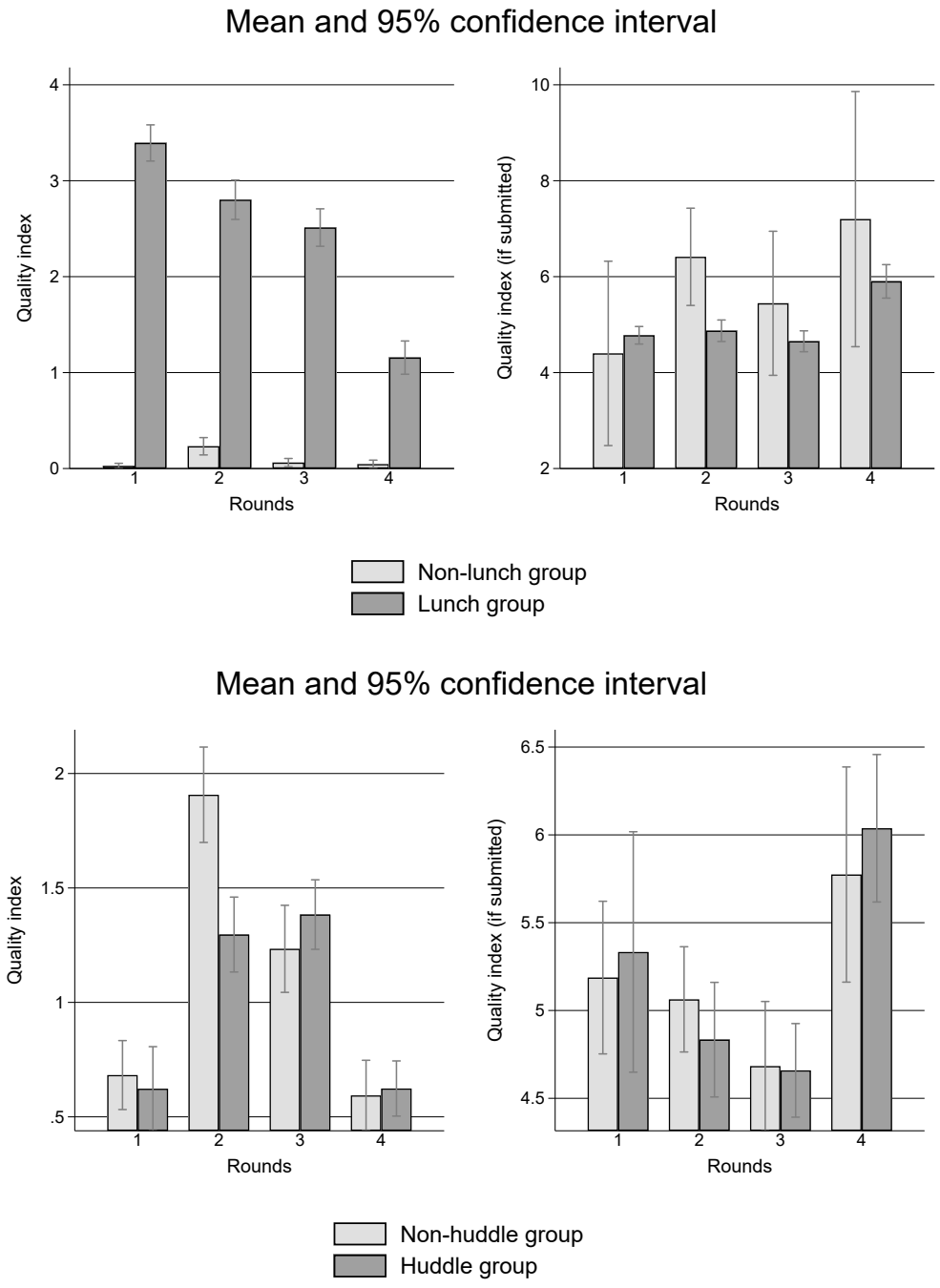


Figure H.17: Quality of submitted suggestions per round

Notes: The upper graphs of Figure H.17 shows the mean and 95% confidence intervall of the quality of suggestions by non-lunch and lunch group per round, whereas the lower graphs show it by non-huddle and huddle group per round. In both left graphs, the quality is set to zero if no suggestion was submitted in the respective round. In both right graphs, the data is restricted to submitted suggestions only.

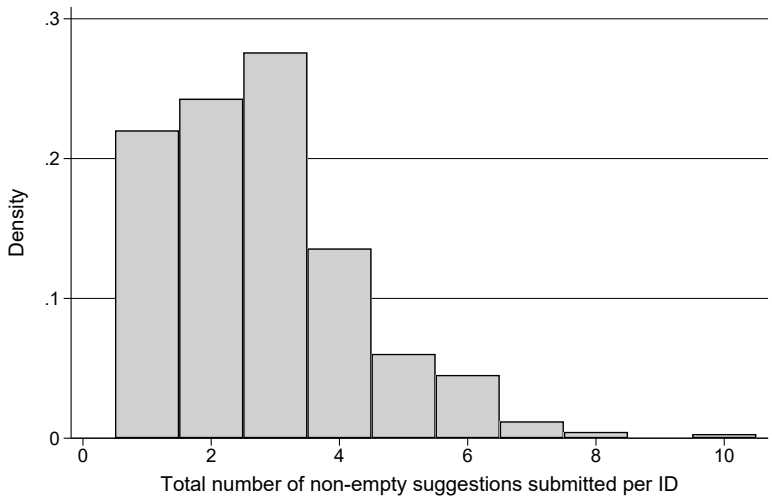


Figure H.18: Total number of submissions per worker

Notes: Figure H.18 shows the distribution of the total number of submissions for workers who submitted at least one suggestion.

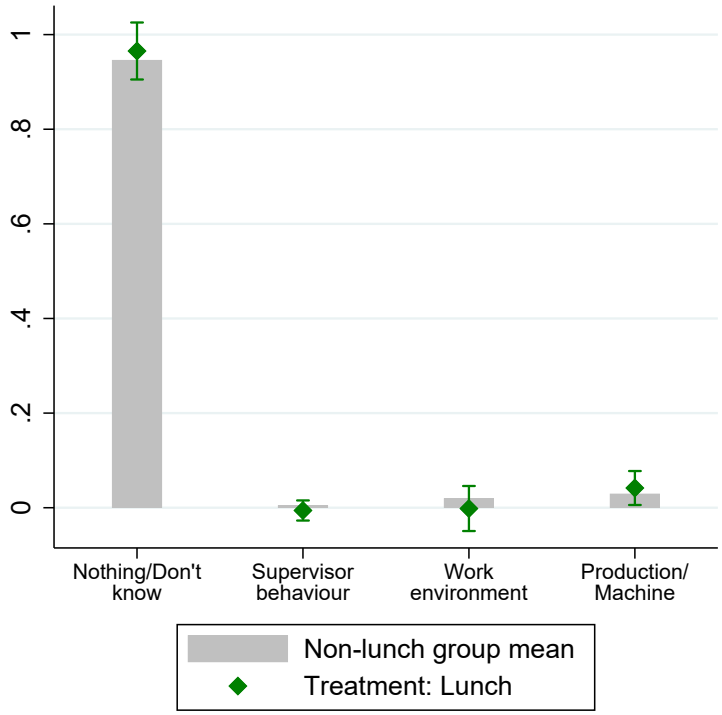


Figure H.19: Perceived implementation of ideas from workers

Notes: Figure H.19 shows the mean and lunch treatment effect for the line operator survey question "Have managers/supervisors implemented any ideas from workers over the past six months?" (N=200). 95% confidence intervals are displayed. Supervisor behavior contains answer options such as reduced shouting or changed working style. Work environment contains answer options such as hand-wash in toilet, water filter, cleaned floor or regular fire alarm checks. Production/Machine contains answer options such as machine replacement or production problem solved.

I Supplementary tables: Integrity of the experiment

TABLE I.11. BALANCE TABLE OF PRIMARY OUTCOME VARIABLES

	Control (1)	Lunch (2)	Huddle (3)	p (1=2)	se (Δ)	p (1=3)	se (Δ)	p (1 = 2 ∪ 3)	se (Δ)	N
Workforce-related outcomes										
Workforce index	-0.01	0.02	0.01	0.03	(0.02)	0.73	(0.02)	0.06	(0.02)	1532
Absenteeism	0.02	0.03	0.02	0.03	(0.00)	0.33	(0.00)	0.02	(0.00)	1532
Punctuality	4.89	4.76	5.04	0.26	(0.24)	0.24	(0.29)	0.47	(0.31)	1531
Working hours	11.04	11.03	11.03	0.63	(0.03)	0.60	(0.04)	0.20	(0.04)	1532
Production-related outcomes										
Production index	0.07		-0.01			0.38	(0.10)			51
Efficiency	0.53		0.53			0.61	(0.01)			51
Alteration rate	0.08		0.08			0.58	(0.00)			51
Reject rate	0.00		0.00			0.35	(0.00)			51
Job satisfaction-related outcomes										
Job satisfaction index	0.14	0.09	0.05	0.86	(0.17)	0.65	(0.20)	0.40	(0.24)	204
Job satisfaction	0.34	0.33	0.31	0.89	(0.13)	0.79	(0.16)	0.91	(0.18)	204
Quit intentions	-0.08	-0.17	-0.23	0.81	(0.18)	0.42	(0.21)	0.28	(0.28)	202
Job attitudes	-0.04	-0.00	0.00	0.58	(0.12)	0.50	(0.13)	0.90	(0.16)	202
Job satisfaction supervisors	0.00		-0.22			0.37	(0.25)			94
Mental health-related outcomes										
Mental health index	0.19	0.05	-0.06	0.54	(0.16)	0.11	(0.20)	0.04	(0.21)	202
Feelings at work	0.16	0.07	-0.01	0.77	(0.17)	0.29	(0.22)	0.16	(0.21)	202
General Anxiety	0.11	-0.01	-0.11	0.61	(0.13)	0.04	(0.14)	0.01	(0.15)	202
Workplace harrassment	0.02	-0.02	0.02	0.90	(0.17)	0.80	(0.20)	0.62	(0.26)	202

Notes: Table I.11 shows balance checks for primary outcome variables across treatment groups. Columns 1 through 3 show sample means in the control group (1), lunch group (2), and the huddle group (3), respectively. Columns 4 through 9 shows p-values and standard deviations of tests for equality of means between the lunch group, the huddle group, and the lunch and huddle group combined, compared to the control groups. P-values are based on standard errors clustered at the worker level for the lunch group, and at the line-level for the huddle group, and the lunch and huddle group combined. The data period for the balance test is three months prior to the intervention start date.

TABLE I.12. BALANCE TABLE OF WORKER CHARACTERISTICS AND SECONDARY OUTCOME VARIABLES

	Operators										Supervisors				
	Control (1)	Lunch (2)	Huddle (3)	p (1=2)	se (Δ)	p (1=3)	se (Δ)	p (1 = 2 ∪ 3)	se (Δ)	N	Control (4)	Huddle (5)	p (4=5)	se (Δ)	N
Panel A: Characteristics															
Female share	0.79	0.77	0.79	0.69	(0.06)	0.62	(0.05)	0.40	(0.06)	204	0.24	0.32	0.41	(0.09)	94
Age	25.94	25.82	26.13	0.70	(0.70)	0.59	(0.66)	0.91	(0.76)	204	30.44	29.77	0.58	(1.20)	94
Experience index	0.01	-0.09	-0.01	0.24	(0.14)	0.83	(0.14)	0.80	(0.18)	204					
Nb factories	0.90	0.75	0.98	0.15	(0.20)	0.41	(0.19)	0.81	(0.19)	204	2.12	2.41	0.58	(0.53)	94
Grade	4.87	4.99	4.85	0.29	(0.19)	0.71	(0.19)	0.73	(0.24)	203					
Tenure factory in years	2.81	2.84	2.61	0.89	(0.48)	0.46	(0.48)	0.94	(0.65)	204	5.66	5.28	0.60	(0.74)	94
Tenure industry in years	4.64	4.88	4.72	0.64	(0.63)	1.00	(0.61)	0.60	(0.76)	204	10.07	9.80	0.79	(1.01)	94
Education	8.53	8.65	8.22	0.58	(0.51)	0.27	(0.50)	0.82	(0.64)	204	6.66	9.70	0.10	(1.83)	94
Processes	4.76	4.86	5.16	0.99	(0.54)	0.39	(0.69)	0.37	(0.66)	140					
Days per line	54.79	55.13	53.49	0.72	(3.39)	0.58	(3.16)	0.71	(3.80)	204	549.40	632.48	0.63	(171.83)	74
Lines	3.59	3.41	3.64	0.25	(0.28)	0.69	(0.26)	0.91	(0.34)	204	1.12	1.11	0.92	(0.07)	94
Supervised workers											18.18	19.16	0.49	(1.40)	94
Difference L-C (in %)											0.19	0.15	0.12	(0.03)	94
Panel B: Channels															
Engagement and support	-0.02	0.12	0.05	0.22	(0.16)	0.74	(0.19)	0.29	(0.20)	204	-0.36	-0.11	0.26	(0.23)	94
Perceived constraints	-0.15	-0.20	-0.39	0.87	(0.19)	0.15	(0.23)	0.13	(0.23)	172	0.20	-0.03	0.20	(0.18)	94
Perceived efficacy	0.08	-0.03	0.03	0.30	(0.16)	0.78	(0.20)	0.66	(0.21)	204	-0.17	0.01	0.35	(0.20)	94
Communication	0.26	0.31	0.16	0.53	(0.19)	0.35	(0.19)	0.64	(0.18)	204	0.15	0.37	0.42	(0.27)	94
Workplace culture	-0.08	-0.21	-0.18	0.33	(0.15)	0.60	(0.17)	0.18	(0.19)	204	0.07	-0.14	0.38	(0.23)	94

Notes: Table I.12 shows balance checks for main indices from survey data from line operators (columns 1 to 10) and line supervisors (columns 11-15) across the treatment groups. Columns 1 through 3 show sample means for line operators in the control group (1), lunch group (2), and the huddle group (3), respectively. Columns 4 and 6 show p-values of tests for equality of means between the lunch and the huddle groups compared to the control groups, respectively. Column 8 shows the corresponding p-values for comparisons between the control group and the lunch and huddle group combined. Columns 11 and 12 show sample means for line supervisors in the control group (11) and the huddle group (12). Column 13 shows p-values of tests for equality of means between the control and the huddle groups. P-values are based on standard errors clustered at the worker level for the lunch group, and at the line-level for the huddle group and the lunch and huddle group combined. The test for joint significance of differences between the control and lunch group is not rejected (p-value of the F-statistic of the omnibus balance test: 0.29), similarly for differences between the control and huddle group (p-value of the F-statistic of the omnibus balance test: 0.58) and between supervisors in the control and huddle group (p-value of the F-statistic of the omnibus balance test: 0.11)

TABLE I.13. EFFECT OF SKILL LEVEL ON LINE MOVEMENTS

	Total number of lines	**Max. days per line**
	(1)	(2)
Number of processes per worker	0.035	-0.73*
	(0.022)	(0.42)
Observations	1177	973
R^2	0.26	0.06
Control Mean	3.63	57.32
Floor FE	YES	YES

Notes: Table I.13 shows an OLS regression of the number of processes each worker can perform (based on skills matrix data) on the total number of lines in column 1 and on the maximum number of days per line in column 2, each based on the intervention period from January to March 2020. All regressions include factory fixed effects. Robust standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE I.14. ATTRITION

	Dep. variable: Attrition		
	(1)	(2)	(3)
Lunch	-0.007		0.001
	(0.752)		(0.969)
Huddle		0.043	0.052
		(0.109)	(0.110)
LunchXHuddle			-0.019
			(0.655)
Observations	1587	1598	1587
R^2	0.00	0.00	0.00
Control Mean	0.27	0.25	0.26
Clusters	794	51	51
p (L=H)			0.12
p (L=LxH)			0.76
p (H=LxH)			0.28

Notes: Table I.14 shows a linear regression of the treatment indicators on a variable equal to one if the worker was not surveyed at endline. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

J Supplementary tables: Main outcomes

TABLE J.15. EFFECTS ON PRIMARY OUTCOME VARIABLES RELATED TO INNOVATION (SAMPLE ONE ONLY)

	Quantity			**Quality**			**Quality (if submitted)**			**Variety**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lunch	0.47***		0.44***	2.28***		2.31***	-1.05**		-1.59**	
	(0.014)		(0.022)	(0.082)		(0.18)	(0.51)		(0.72)	
Huddle		-0.0069	-0.017		-0.071	-0.13		-0.067	-0.39	0.063
		(0.017)	(0.011)		(0.12)	(0.085)		(0.20)	(1.09)	(0.056)
LunchXHuddle			-0.012			-0.051			0.38	
			(0.029)			(0.21)			(1.08)	
Observations	5672	4577	4577	5672	4577	4577	1222	1059	1059	439
R^2	0.36	0.07	0.33	0.26	0.06	0.25	0.06	0.09	0.10	0.05
Control Mean	0.02	0.23	0.16	0.10	1.27	0.87	6.60	5.44	5.46	1.16
Round FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
p (L=H)			0.00			0.00			0.15	
p (L=LxH)			0.00			0.00			0.24	
p (H=LxH)			0.89			0.77			0.72	

Notes: All regression exclude observations from workers in randomisation sample two, i.e. workers without friends at baseline. In columns 1, 4 and 7 standard errors are clustered at the pair level. In all other columns, standard errors are clustered at the line level, and observations prior to the first huddle (14th of January 2020) are excluded. Columns 7 to 10 only include submitted suggestions. Column 10 further restricts observations to individuals who submitted more than once (which are only some of the lunch workers). The quality index (column 4 to 9) consists of six components, out of which two are criteria to evaluate the usefulness of submitted ideas for workers, namely impact on worker well-being and implementation level, and four are criteria to evaluate the usefulness of ideas for the firm, namely financial resources, impact on efficiency, implementation time and risk, and level of effects. * p<0.1, ** p<0.05, *** p<0.01

TABLE J.16. EFFECTS ON PRIMARY OUTCOME VARIABLES RELATED TO FIRM PERFORMANCE AND WORKER WELL-BEING (SAMPLE ONE ONLY)

	Workforce index			**Production index**	**Job satisfaction index**			**Mental health index**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lunch	-0.007		-0.003		0.212**		0.309*	0.021		0.072
	(0.716)		(0.901)		(0.024)		(0.099)	(0.862)		(0.601)
Huddle		0.009	0.022	-0.162		0.117	0.187		0.172	0.225
		(0.667)	(0.427)	(0.182)		(0.194)	(0.197)		(0.282)	(0.232)
LunchXHuddle			-0.023				-0.203			-0.145
			(0.549)				(0.330)			(0.510)
Observations	107531	82375	82375	459	195	195	195	195	195	195
R^2	0.10	0.09	0.09	0.50	0.21	0.19	0.23	0.13	0.15	0.18
Control Mean	-0.14	-0.10	-0.10	0.15	-0.23	-0.17	-0.14	-0.13	-0.17	-0.12
Clusters	707	703	703	51	182	51	51	182	51	51
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
p (L=H)			0.38				0.36			0.41
p (L=LxH)			0.74				0.17			0.51
p (H=LxH)			0.46				0.25			0.31

Notes: All regression exclude observations from workers in randomisation sample two, i.e. workers without friends at baseline. Standard errors are clustered at the pair level in columns 1, 5, and 8, and at the line level otherwise. Time fixed effects are included in regressions on the workforce index (in days) and on the production index (in weeks). Huddle exposure is truncated at one (at least one huddle). The cumulative exposure per worker is used in columns 1, 2 and 3, the average cumulative exposure per line in column 4, and the maximum exposure per worker in columns 5 to 10. Indices are constructed using the Anderson (2008) method. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE J.17. LATE (IV) EFFECTS ON PRIMARY OUTCOMES (INNOVATION)

	Quantity			**Ideas**			**Ideas (if submitted)**			**Variety**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lunch	0.61***		0.56***	0.26***		0.24***	-0.13*		-0.19*	
	(0.018)		(0.084)	(0.011)		(0.064)	(0.081)		(0.10)	
Huddle		-0.065	-0.063		-0.027	-0.047		0.0021	-0.072	0.30
		(0.078)	(0.046)		(0.055)	(0.038)		(0.033)	(0.18)	(0.26)
LunchXHuddle			-0.082			0.0086			0.080	
			(0.12)			(0.089)			(0.18)	
Observations	6352	4957	4957	6710	5179	5179	1857	1294	1294	544
R^2	0.32	0.06	0.27	0.18	0.04	0.16	0.06	0.07	0.07	-0.07
Control Mean	0.03	0.40	0.25	0.01	0.15	0.09	0.67	0.53	0.53	1.20
Round FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Kleib.-P. LM	0	0	0	0	0	0				0
Kleib.-P. Wald F	8388.87	61.48	13.75	10235.44	67.83	15.64				25.23
p (L=H)			0.00			0.00			0.43	
p (L=LxH)			0.00			0.13			0.29	
p (H=LxH)			0.89			0.60			0.67	

Notes: In columns 1, 4 and 7 standard errors are clustered at the pair level. In all other columns, standard errors are clustered at the line level, and observations prior to the first huddle (14th of January 2020) are excluded. Columns 7 to 10 only include submitted suggestions. Column 10 further restricts observations to individuals who submitted more than once (which are only some of the lunch workers). The quality index (column 4 to 9) consists of six components, out of which two are criteria to evaluate the usefulness of submitted ideas for workers, namely impact on worker well-being and implementation level, and four are criteria to evaluate the usefulness of ideas for the firm, namely financial resources, impact on efficiency, implementation time and risk, and level of effects (separate regressions shown in Table J.21). * p<0.1, ** p<0.05, *** p<0.01

TABLE J.18. LATE (IV) EFFECTS ON PRIMARY OUTCOMES (FIRM PERFORMANCE AND WORKER WELL-BEING)

	Workforce index			**Production index**	**Job satisfaction index**			**Mental health index**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Lunch	-0.006		0.045		0.196**		0.924	0.016		0.625
	(0.733)		(0.594)		(0.032)		(0.248)	(0.889)		(0.387)
Huddle		0.020	0.074	-0.315		0.612	1.237		0.969	1.521
		(0.781)	(0.447)	(0.170)		(0.235)	(0.280)		(0.310)	(0.294)
LunchXHuddle			-0.108				-1.212			-1.065
			(0.478)				(0.314)			(0.378)
Observations	119588	92078	91580	459	195	195	195	195	195	195
R^2	0.05	0.05	0.05	0.50	0.20	0.00	-0.13	0.13	-0.12	-0.17
Control Mean	-0.18	-0.13	-0.11	0.15	-0.23	-0.14	-0.15	-0.13	-0.21	-0.12
Clusters	792	55	55	51	182	51	51	182	51	51
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Kleib.-P. LM	0	0	0	0		0	0		0	0
Kleib.-P. Wald F	27004.69	67.09	17.88	233.81		7.55	1.22		5.95	1.14
p (L=H)			0.72				0.59			0.38
p (L=LxH)			0.51				0.29			0.38
p (H=LxH)			0.43				0.29			0.30

Notes: Standard errors are clustered at the pair level in columns 1, 5, and 8, and at the line level otherwise. Time fixed effects are included in regressions on the workforce index (in days) and on the production index (in weeks). Huddle exposure is truncated at one (at least one huddle). The cumulative exposure per worker is used in columns 1, 2 and 3, the average cumulative exposure per line in column 4, and the maximum exposure per worker in columns 5 to 10. Indices are constructed using the Anderson (2008) method. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE J.19. EFFECTS ON SUB-COMPONENTS (FIRM PERFORMANCE AND WORKER WELL-BEING)

Panel A: ITT Effects of lunch treatment

	Workforce index				Production index			Job satisfaction index			Mental health index		
	(1) Absenteeism	(2) Punctuality	(3) Working time	(4) Exits*	(5) Efficiency	(6) Alterations	(7) Rejects	(8) Job satisfaction	(9) Quit intentions	(10) Job attitudes	(11) Feelings at work	(12) General anxiety	(13) Harrassment
Lunch	0.001	0.015	0.004	0.001				0.130	0.033	0.365**	-0.052	0.054	0.208*
	(0.002)	(0.224)	(0.025)	(0.015)				(0.128)	(0.082)	(0.151)	(0.104)	(0.140)	(0.121)
	{1.000}	{1.000}	{1.000}	{1.000}				{0.454}	{0.856}	{0.052}	{0.872}	{0.872}	{0.360}
Observations	119588	91591	91747	1450	447	459	459	195	195	195	195	195	194
Clusters	792	788	788	784	51	51	51	182	182	182	182	182	181
R^2	0.02	0.11	0.21	0.03	0.47	0.52	0.25	0.14	0.07	0.27	0.03	0.28	0.07
Control mean	0.03	-3.96	-11.80	0.09	0.59	-0.08	-0.00	-0.44	0.11	-0.16	-0.09	-0.09	-0.09
Time FE	YES	YES	YES	NO	YES	YES	YES	NO	NO	NO	NO	NO	NO
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Panel B: ITT Effects of lunch and huddle treatments

	Workforce index				Production index			Job satisfaction index			Mental health index		
	(1) Absenteeism	(2) Punctuality	(3) Working time	(4) Exits*	(5) Efficiency	(6) Alterations	(7) Rejects	(8) Job satisfaction	(9) Quit intentions	(10) Job attitudes	(11) Feelings at work	(12) General anxiety	(13) Harrassment
Lunch	0.003	0.041	-0.001	-0.001				0.162	0.119	0.471**	-0.070	0.185	0.093
	(0.002)	(0.262)	(0.020)	(0.024)				(0.148)	(0.170)	(0.183)	(0.122)	(0.192)	(0.166)
	{1.000}	{1.000}	{1.000}	{1.000}				{0.384}	{0.480}	{0.042}	{1.000}	{1.000}	{1.000}
Huddle	0.002	-0.371	0.075*	-0.001	-0.005	-0.007*	-0.000	0.435**	0.139	0.053	0.158	0.185	-0.112
	(0.003)	(0.298)	(0.038)	(0.022)	(0.014)	(0.004)	(0.000)	(0.205)	(0.152)	(0.210)	(0.162)	(0.204)	(0.176)
	{0.776}	{0.488}	{0.278}	{0.933}	{0.920}	{0.194}	{0.920}	{0.135}	{0.571}	{1.000}	{1.000}	{1.000}	{1.000}
LunchXHuddle	-0.005	-0.033	0.008	0.005				-0.124	-0.181	-0.207	-0.009	-0.264	0.223
	(0.004)	(0.360)	(0.042)	(0.031)				(0.210)	(0.187)	(0.311)	(0.257)	(0.240)	(0.258)
	{1.000}	{1.000}	{1.000}	{1.000}				{1.000}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}
Observations	119588	91591	91747	1450	447	459	459	195	195	195	195	195	194
Clusters	51	51	51	51	51	51	51	49	49	49	49	49	49
R^2	0.02	0.12	0.21	0.03	0.47	0.52	0.25	0.28	0.11	0.28	0.06	0.31	0.11
Control mean	0.03	-3.82	-11.90	0.09	0.59	-0.08	-0.00	-0.61	0.03	-0.21	-0.17	-0.19	-0.07
Time FE	YES	YES	YES	NO	YES	YES	YES	NO	NO	NO	NO	NO	NO
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
p (L=H)	0.58	0.18	0.09	0.98				0.16	0.86	0.08	0.15	1.00	0.29
q (L=H)	0.63	0.55	0.55	0.97				0.32	0.40	0.32	0.77	0.77	0.77
p (L=LxH)	0.20	0.89	0.86	0.92				0.39	0.39	0.13	0.86	0.27	0.74
q (L=LxH)	1.00	1.00	1.00	1.00				0.65	0.65	0.65	1.00	1.00	1.00
p (H=LxH)	0.34	0.57	0.36	0.90				0.14	0.33	0.60	0.66	0.26	0.42
q (H=LxH)	1.00	1.00	1.00	1.00				0.73	0.73	0.73	1.00	1.00	1.00

Notes: Panel A shows ITT effects of the lunch treatment, and panel B of the lunch and the huddle treatments. Effects on the production index can solely be displayed for the huddle treatment due to unavailability of production data at the worker-level. Variables *Exits* (column 4), denoted with a star, is collapsed at the individual worker-level and not included in the workforce index. Standard errors are clustered at the pair level in panel A and at the line level in panel B. False discovery rate (FDR) adjusted p-values, also known as q-values, were used to correct for multiple hypothesis testing. They are shown in curly brackets. These were calculated following the method of Benjamini et al. (2006). Standard errors are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE J.20. QUALITY THRESHOLD

	Quality			**Quality (if submitted)**		
	(1)	(2)	(3)	(4)	(5)	(6)
Lunch	0.21***		0.22***	-0.11		-0.19*
	(0.0093)		(0.020)	(0.086)		(0.099)
Huddle		-0.0064	-0.014		0.0021	-0.071
		(0.013)	(0.010)		(0.033)	(0.18)
LunchXHuddle			0.0017			0.079
			(0.023)			(0.18)
Observations	6352	5179	5179	1499	1294	1294
R^2	0.14	0.04	0.14	0.06	0.07	0.07
Control Mean	0.01	0.14	0.10	0.67	0.53	0.53
Round FE	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES
p (L=H)			0.00			0.44
p (L=LxH)			0.00			0.29
p (H=LxH)			0.59			0.67

Notes: Quality is an indicator equal to one if the suggestion is an idea (type five and six). In columns 1 and 4, standard errors are clustered at the pair level. In all other columns, standard errors are clustered at the line level, and observations prior to the first huddle (14th of January 2020) are excluded. Columns 7 to 10 only include submitted suggestions. * p<0.1, ** p<0.05, *** p<0.01

TABLE J.21. EFFECTS ON QUALITY INDEX COMPONENTS

	Worker comp.			**Firm comp.**		
	(1)	(2)	(3)	(4)	(5)	(6)
Lunch	-0.57		-0.68	-1.27**		-1.39*
	(0.41)		(0.61)	(0.50)		(0.73)
Huddle		-0.096	-0.11		0.011	0.27
		(0.18)	(0.90)		(0.18)	(1.21)
LunchXHuddle			0.034			-0.23
			(0.88)			(1.19)
Observations	1857	1294	1294	1857	1294	1294
R^2	0.06	0.08	0.08	0.05	0.06	0.07
Control Mean	6.00	5.44	5.45	6.46	5.24	5.27
Round FE	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES
p (L=H)			0.39			0.09
p (L=LxH)			0.60			0.51
p (H=LxH)			0.93			0.84

Notes: Table J.21 shows individual regressions for worker and firm components fo the quality index. Worker components include the two criteria impact on worker well-being and implementation level. Firm components include the four criteria financial resources, impact on efficiency, implementation time and risk, and level of effects (direct or indirect). * p<0.1, ** p<0.05, *** p<0.01

TABLE J.22. EFFECTS ON DIFFERENT SPECIFICATIONS OF VARIETY

	Categorization			**Latent Semantic Analysis**
	(1) Category	(2) Subcategory	(3) Suggestion	(4) Suggestion text
Huddle	0.033	0.059	0.063	0.011
	(0.062)	(0.061)	(0.053)	(0.015)
Observations	544	544	544	543
R^2	0.03	0.05	0.05	0.03
Control Mean	0.87	0.94	1.13	0.89
Round FE	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES

Notes: Variety in columns 1 to 3 is composed of the number of new categories, subcategories or suggestions resp. of each suggestion paper in comparison with the ones mentioned in any prior paper by the same worker. Variety in column 4 is composed of the percentage of new content based on a latent semantic analysis approach which compares word vectors of each suggestion paper with the ones of any prior paper by the same worker.

TABLE J.23. EFFECTS ON THE CONTENT OF SUGGESTIONS

Panel (a): ITT Effects of lunch treatment

	Supervisor skills		Attendance bonus/Leave		Work environment				Production				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Behaviour and tasks	Gender	Leave	Attend. bonus	WE (behavioral/facilities)	WE (workplace)	Salary	Health and safety	Manpower	Machine	Motion/waiting time	Efficiency	Production layout
Lunch	-0.057	0.004	-0.171**	0.067	-0.004	0.002	0.035***	0.022***	0.055*	0.003	0.026***	0.018**	0.000
	(0.059)	(0.005)	(0.070)	(0.049)	(0.045)	(0.065)	(0.009)	(0.007)	(0.032)	(0.027)	(0.008)	(0.008)	(0.001)
	{0.676}	{0.676}	{0.032}	{0.092}	{0.956}	{0.956}	{0.001}	{0.002}	{0.103}	{0.569}	{0.007}	{0.049}	{0.569}
Observations	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015
Clusters	356	356	356	356	356	356	356	356	356	356	356	356	356
R^2	0.05	0.05	0.06	0.02	0.01	0.07	0.01	0.01	0.06	0.05	0.02	0.03	0.01
Control mean	0.19	0.00	0.32	0.11	0.08	0.22	0.00	0.00	0.03	0.05	0.00	0.00	0.00
Time FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Panel (b): ITT Effects of huddle treatment

	Supervisor skills		Attendance bonus/Leave		Work environment				Production				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Behaviour and tasks	Gender	Leave	Attend. bonus	WE (behavioral/facilities)	WE (workplace)	Salary	Health and safety	Manpower	Machine	Motion/waiting time	Efficiency	Production layout
Huddle	-0.043**	-0.003	-0.006	0.007	-0.003	0.045	-0.013	0.003	0.015	0.011	-0.009	-0.002	-0.002
	(0.020)	(0.005)	(0.023)	(0.021)	(0.016)	(0.027)	(0.011)	(0.009)	(0.023)	(0.013)	(0.006)	(0.008)	(0.002)
	{0.074}	{0.441}	{1.000}	{1.000}	{0.927}	{0.751}	{0.751}	{0.927}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}
Observations	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015
Clusters	51	51	51	51	51	51	51	51	51	51	51	51	51
R^2	0.05	0.05	0.05	0.02	0.02	0.07	0.01	0.01	0.06	0.05	0.02	0.03	0.01
Control mean	0.19	0.01	0.16	0.17	0.07	0.16	0.04	0.02	0.08	0.05	0.02	0.02	0.00
Time FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Panel (c): ITT Effects of lunch and huddle treatment

	Supervisor skills		Attendance bonus/Leave		Work environment				Production				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Behaviour and tasks	Gender	Leave	Attend. bonus	WE (behavioral/facilities)	WE (workplace)	Salary	Health and safety	Manpower	Machine	Motion/waiting time	Efficiency	Production layout
Lunch	-0.042	-0.001	-0.141**	0.058	-0.043	0.022	0.046***	0.024*	0.011	0.015	0.031**	0.019*	0.001
	(0.078)	(0.005)	(0.069)	(0.062)	(0.080)	(0.058)	(0.011)	(0.013)	(0.043)	(0.024)	(0.013)	(0.011)	(0.002)
	{1.000}	{1.000}	{0.103}	{0.220}	{0.553}	{0.553}	{0.001}	{0.096}	{0.913}	{0.689}	{0.105}	{0.203}	{0.689}
Huddle	0.015	-0.021**	0.099	-0.024	-0.132*	0.092	0.007	0.001	-0.099	0.064	-0.004	0.003	0.000
	(0.118)	(0.009)	(0.145)	(0.097)	(0.077)	(0.128)	(0.009)	(0.013)	(0.060)	(0.069)	(0.011)	(0.014)	(0.002)
	{0.818}	{0.038}	{1.000}	{1.000}	{0.609}	{0.915}	{0.915}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}
LunchXHuddle	-0.059	0.019*	-0.103	0.030	0.133	-0.050	-0.021	0.002	0.118	-0.055	-0.006	-0.006	-0.003
	(0.117)	(0.010)	(0.149)	(0.101)	(0.081)	(0.128)	(0.017)	(0.017)	(0.073)	(0.070)	(0.013)	(0.017)	(0.003)
	{0.447}	{0.134}	{1.000}	{1.000}	{0.717}	{0.812}	{0.717}	{0.812}	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}
Observations	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015	1015
Clusters	51	51	51	51	51	51	51	51	51	51	51	51	51
R^2	0.05	0.05	0.07	0.03	0.02	0.07	0.02	0.01	0.07	0.05	0.03	0.03	0.01
Control mean	0.16	0.00	0.32	0.12	0.12	0.20	0.00	0.00	0.04	0.04	0.00	0.00	0.00
Time FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
p (L=H)	0.51	0.01	0.06	0.23	0.00	0.56	0.00	0.02	0.01	0.47	0.00	0.10	0.44
q (L=H)	0.34	0.01	0.15	0.15	0.00	0.16	0.00	0.01	0.03	0.23	0.00	0.11	0.23
p (L=LxH)	0.93	0.15	0.85	0.86	0.27	0.66	0.02	0.42	0.34	0.39	0.14	0.35	0.41
q (L=LxH)	0.87	0.41	1.00	1.00	0.69	0.99	0.07	0.74	0.69	0.69	0.69	0.69	0.69
p (H=LxH)	0.75	0.03	0.49	0.78	0.10	0.58	0.26	0.96	0.11	0.39	0.90	0.78	0.47
q (H=LxH)	0.61	0.06	1.00	1.00	0.64	1.00	0.65	1.00	1.00	1.00	1.00	1.00	1.00

Notes: Panel (a) shows lunch treatment effects, panel (b) huddle treatment effects, and panel (c) lunch and huddle treatment effects on the content of suggestions. Regressions in panel (b) and (c) exclude all suggestions prior to the first day a huddle took place (i.e. 14th of January 2020). Categories are only defined for type four and five (problems and ideas).

K Supplementary tables: Mechanisms

TABLE K.24. LEARNING OVER TIME

	Quantity		**Quality**		**Variety**
	(1)	(2)	(3)	(4)	(5)
Round2	0.030***	0.030***	0.20***	0.20***	
	(0.0093)	(0.0092)	(0.061)	(0.061)	
Round3	0.0048	0.0050	0.033	0.033	-0.11*
	(0.0052)	(0.0052)	(0.033)	(0.033)	(0.061)
Round4		-0.0000070		0.018	0.11
		(0.0061)		(0.045)	(0.070)
LunchXRound1	0.70***	0.71***	3.23***	3.71***	
	(0.020)	(0.034)	(0.12)	(0.24)	
LunchXRound2	0.51***	0.61***	2.36***	3.17***	
	(0.023)	(0.038)	(0.15)	(0.26)	
LunchXRound3	0.53***	0.51***	2.52***	2.25***	
	(0.021)	(0.040)	(0.13)	(0.23)	
LunchXRound4		0.56***		3.24***	
		(0.037)		(0.25)	
Observations	4401	4104	4401	4104	543
R^2	0.44	0.51	0.31	0.40	0.03
Control Mean	1.00	0.03	1.60	0.03	1.40
Floor FE	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES
p (LunchXRound1=LunchXRound2)	0.00	0.02	0.00	0.09	
p (LunchXRound2=LunchXRound3)	0.45	0.04	0.36	0.00	
p (LunchXRound3=LunchXRound4)		0.29		0.00	
p (Round3=Round4)					0.00

Notes: Quantity is the number of suggestions per worker per round. Quality is the quality index on a scale from one to ten with non-submissions coded as zero. Variety is the number of categories per worker per round, which have not been mentioned by that worker in any previous rounds. A round is defined as the time between a lunch of sub-group one and the day before the subsequent lunch of sub-group one. Per round, each of the four sub-groups receive one lunch. Columns 1 and 3 include observations of treated workers who were assigned to receive three lunches. Columns 2 and 4 only include observations of treated workers who were assigned to receive four lunches. Column 5 excludes observations from workers who did not submit more than one suggestion, and hence, only includes the lunch group. It further excludes observations from the first round, as variety can only be measured if at least one suggestion per worker has already been submitted. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.25. EFFECTS ON SECONDARY OUTCOME COMPONENTS

	Operators								Supervisors	
	(1) Control μ	(2) Lunch	(3) Control μ	(4) Huddle	(5) Control μ	(6) Lunch	(7) Huddle	(8) Both	(9) Control μ	(10) Huddle
Futility Index	-0.04	0.35***	-0.01	0.08	-0.03	0.29	-0.01	0.11	-0.05	0.02
		(0.12)		(0.13)		(0.19)	(0.17)	(0.17)		(0.22)
Predicted number of own submissions	0.03	0.05*	0.06	0.03	0.12	0.02	-0.00	0.06		
		(0.03)		(0.02)		(0.03)	(0.04)	(0.07)		
Belief about winning	-0.04	0.38***	-0.01	0.09	-0.01	0.36**	0.04	0.03		
		(0.11)		(0.12)		(0.16)	(0.15)	(0.16)		
Belief about implementation of winner's idea	-0.05	0.22	-0.05	0.08	-0.13	0.16	-0.00	0.11	-0.04	0.12
		(0.16)		(0.15)		(0.23)	(0.21)	(0.23)		(0.22)
Belief about implementation of own idea	-0.09	0.40***	-0.10	0.06	-0.22	0.38*	-0.03	0.08		
		(0.15)		(0.15)		(0.22)	(0.20)	(0.22)		
Feeling incentivized to submit idea (not in index, N=1164)	0.28	0.11*	0.31	-0.03	0.35	0.12	-0.04	-0.03		
		(0.06)		(0.06)		(0.08)	(0.09)	(0.11)		
Likelihood of prize on own line									-0.09	0.33*
										(0.17)
Satisfaction with award of prizes									-0.01	0.15
										(0.19)
Workers as biggest beneficiary of suggestion box									-0.05	-0.11
										(0.21)
Motivation of typical worker in factory vs line									-0.00	-0.09
										(0.19)
Concentration of typical worker in factory vs line									0.06	-0.28
										(0.20)
Time Index	-0.11	0.09	-0.04	0.20	-0.14	0.06	0.15	0.05	-0.02	-0.05
		(0.12)		(0.14)		(0.16)	(0.20)	(0.26)		(0.24)
Time as perceived constraint	-0.02	-0.02	0.01	0.08	-0.04	-0.05	0.05	0.05	0.06	-0.08
		(0.16)		(0.13)		(0.25)	(0.24)	(0.36)		(0.27)
Frequency thinking about problems	-0.06	0.07	-0.03	0.04	-0.08	0.20	0.17	-0.27	-0.06	-0.01
		(0.14)		(0.14)		(0.17)	(0.20)	(0.31)		(0.16)
Frequency thinking about solutions	-0.09	0.12	0.00	0.03	-0.03	0.18	0.08	-0.14	-0.07	-0.01
		(0.12)		(0.11)		(0.17)	(0.18)	(0.26)		(0.19)
Thinking about ideas over lunch	-0.05	0.04	-0.05	0.24*	-0.04	-0.26	-0.09	0.59*		
		(0.15)		(0.14)		(0.22)	(0.22)	(0.34)		
Frequency discussing ideas outside of factory	-0.10	0.08	-0.02	0.10	-0.12	0.19	0.21	-0.25		
		(0.13)		(0.12)		(0.17)	(0.20)	(0.28)		
Knowledge Index	-0.02	0.04	-0.00	-0.00	0.02	-0.05	-0.09	0.16	-0.06	0.15
		(0.15)		(0.12)		(0.17)	(0.21)	(0.27)		(0.23)
Knowledge as perceived constraint	0.01	-0.06	0.05	0.14	0.08	-0.22	-0.05	0.37	-0.02	0.10
		(0.13)		(0.12)		(0.14)	(0.18)	(0.24)		(0.27)
Easiness to think about problems	-0.05	-0.01	-0.03	-0.06	-0.04	-0.13	-0.19	0.22		
		(0.13)		(0.11)		(0.15)	(0.18)	(0.24)		
Easiness to think about solutions	-0.03	0.03	-0.05	0.11	-0.04	0.12	0.22	-0.25		
		(0.13)		(0.12)		(0.17)	(0.20)	(0.28)		
Ability to develop suggestions factory vs line worker									-0.07	0.08
										(0.24)
Number of Observations	200		200		200				97	

Notes: Standard errors are clustered at the pair level for regressions in columns 1 and 2, and at the line level for all other regressions. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.26. EFFECTS ON SECONDARY OUTCOME COMPONENTS

	Operators								Supervisors	
	(1) Control μ	(2) Lunch	(3) Control μ	(4) Huddle	(5) Control μ	(6) Lunch	(7) Huddle	(8) Both	(9) Control μ	(10) Huddle
Security Index	0.01	0.01	0.00	-0.11	0.00	-0.08	-0.22	0.21	0.02	-0.07
		(0.10)		(0.12)		(0.18)	(0.16)	(0.21)		(0.18)
Subindex: Voice climate	0.03	0.06	0.03	-0.20*	0.06	-0.07	-0.34**	0.26	0.08	-0.02
		(0.11)		(0.11)		(0.17)	(0.16)	(0.23)		(0.17)
Sharing of personal issues with superior	0.02	-0.12	0.04	-0.14	-0.04	0.04	0.05	-0.33		
		(0.12)		(0.15)		(0.19)	(0.17)	(0.20)		
Operators sharing issues with superior	0.04	0.07	-0.01	-0.05	-0.02	0.19	0.07	-0.29	-0.07	-0.01
		(0.14)		(0.11)		(0.16)	(0.17)	(0.21)		(0.20)
Importance of sharing personal issues	0.01	0.05	0.07	0.05	0.10	-0.02	-0.01	0.11	-0.04	0.24
		(0.12)		(0.09)		(0.17)	(0.16)	(0.28)		(0.25)
Ability to be yourself at work	-0.05	0.23*	-0.01	-0.04	-0.13	0.28*	-0.01	-0.11		
		(0.12)		(0.11)		(0.15)	(0.13)	(0.22)		
Ability to openly state opinions	-0.06	0.27**	0.06	0.00	0.00	0.23	-0.08	0.06	0.04	-0.03
		(0.12)		(0.12)		(0.21)	(0.22)	(0.34)		(0.14)
Fear of doing something wrong	0.09	-0.33**	0.09	-0.31*	0.22	-0.53**	-0.52**	0.49	0.03	-0.13
		(0.14)		(0.18)		(0.20)	(0.25)	(0.34)		(0.23)
Fear of admitting an error	0.02	0.05	-0.01	0.01	0.06	0.13	0.10	-0.16	0.06	0.08
		(0.12)		(0.14)		(0.13)	(0.15)	(0.19)		(0.18)
Misreporting to supervisor	-0.08	0.39***	-0.01	-0.14	-0.09	0.39**	-0.19	0.05	-0.07	0.08
		(0.15)		(0.15)		(0.18)	(0.21)	(0.26)		(0.19)
Degree of trust between supervisor-worker (telling secret)	0.05	-0.03	-0.04	-0.17	0.05	-0.12	-0.27	0.16	0.05	0.08
		(0.13)		(0.15)		(0.13)	(0.19)	(0.25)		(0.19)
Degree of trust to workers (lending money)									0.07	0.08
										(0.15)
Feeling valued if suggestion made	-0.05	-0.03	0.03	-0.07	0.02	-0.11	-0.12	0.12		
		(0.16)		(0.15)		(0.19)	(0.19)	(0.29)		
Feeling comfortable to submit idea	0.29	0.16***	0.33	-0.08	0.37	0.15**	-0.11	0.03		
		(0.06)		(0.06)		(0.07)	(0.08)	(0.12)		
Conflict between workers and managers	0.04	-0.13	0.05	-0.08	0.11	-0.30*	-0.26	0.38	0.09	-0.29
		(0.14)		(0.18)		(0.15)	(0.22)	(0.24)		(0.18)
Feeling responsible for changes on production line	-0.01	0.14	-0.05	-0.03	-0.09	0.15	-0.01	-0.04		
		(0.09)		(0.08)		(0.15)	(0.14)	(0.24)		
Operator suggestion culture									-0.03	-0.10
										(0.24)
Influence over workers									0.02	0.12
										(0.23)
Importance group goal setting vs supervisor goal setting									0.03	0.02
										(0.19)
Subindex: Leadership efficacy	-0.00	-0.11	-0.01	0.02	-0.04	-0.07	0.08	-0.10		
		(0.12)		(0.11)		(0.14)	(0.14)	(0.18)		
Caring culture (index)	0.04	-0.07	-0.04	0.06	-0.05	-0.14	0.00	0.11		
		(0.10)		(0.11)		(0.13)	(0.16)	(0.21)		
Feeling listened to	0.00	-0.04	0.02	0.04	-0.06	-0.10	-0.02	0.12		
		(0.08)		(0.07)		(0.12)	(0.10)	(0.15)		
Feeling treated with respect	0.04	-0.02	-0.03	0.03	-0.02	-0.10	-0.06	0.16		
		(0.08)		(0.07)		(0.09)	(0.09)	(0.11)		
Feeling appreciated and recognized	0.02	-0.03	-0.01	-0.03	0.02	-0.21**	-0.22*	0.34**		
		(0.09)		(0.08)		(0.10)	(0.12)	(0.15)		
Support from others	0.07	-0.12	-0.02	0.07	-0.02	-0.16	0.09	0.00		
		(0.14)		(0.16)		(0.15)	(0.22)	(0.27)		
Effort of supervisor to get to know worker	-0.01	0.07	-0.05	0.01	-0.07	0.08	0.03	-0.06		
		(0.13)		(0.14)		(0.18)	(0.21)	(0.28)		
Efforts to get to know each other	0.00	-0.07	-0.05	0.11	-0.02	-0.09	0.09	0.06		
		(0.12)		(0.13)		(0.14)	(0.20)	(0.26)		
Supervisor abilities (index)	-0.00	-0.15	0.03	0.02	-0.03	-0.08	0.10	-0.16	0.00	-0.21
		(0.13)		(0.11)		(0.15)	(0.17)	(0.25)		(0.20)
Ability of own supervisor to encourage operators	-0.04	0.03	-0.03	-0.06	-0.13	0.07	0.05	-0.20	-0.03	-0.06
		(0.14)		(0.13)		(0.15)	(0.22)	(0.31)		(0.23)
Ability of typical supervisor to encourage operators	-0.09	0.15	-0.01	-0.05	-0.10	0.36*	0.21	-0.53	-0.06	0.29*
		(0.15)		(0.13)		(0.21)	(0.21)	(0.34)		(0.15)
Delta encouragements between own and typical supervisor	0.07	-0.14	-0.01	0.00	-0.01	-0.34*	-0.19	0.41	0.04	-0.30
		(0.13)		(0.11)		(0.17)	(0.23)	(0.31)		(0.20)
Overall ability of own supervisor	-0.09	0.01	0.12	-0.05	-0.03	0.09	0.00	-0.14	-0.02	0.10
		(0.12)		(0.11)		(0.15)	(0.15)	(0.21)		(0.22)
Overall ability of typical supervisor	-0.15	0.22*	0.08	-0.12	-0.10	0.28	-0.08	-0.13	-0.01	-0.07
		(0.12)		(0.12)		(0.19)	(0.17)	(0.23)		(0.20)
Delta overall ability between own and typical supervisor	0.07	-0.22*	0.03	0.08	0.07	-0.22	0.07	0.03	0.01	-0.18
		(0.13)		(0.10)		(0.16)	(0.17)	(0.20)		(0.20)
Decision-making authority as perceived constraint	-0.01	-0.02	0.00	-0.13	0.00	0.00	-0.12	-0.02	-0.04	0.14
		(0.14)		(0.13)		(0.22)	(0.26)	(0.35)		(0.24)
Number of Observations	200		200		200				97	

Notes: Standard errors are clustered at the pair level for regressions in columns 1 and 2, and at the line level for all other regressions. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.27. COMMUNICATION PRACTICES

	Operators								Supervisors	
	(1) Control μ	(2) Lunch	(3) Control μ	(4) Huddle	(5) Control μ	(6) Lunch	(7) Huddle	(8) Both	(9) Control μ	(10) Huddle
Communication Index	-0.02	-0.11	0.07	-0.03	-0.01	-0.05	0.08	-0.15	-0.06	0.00
		(0.10)		(0.07)		(0.11)	(0.08)	(0.16)		(0.12)
Subindex: General communication	-0.01	0.05	0.04	-0.00	0.07	-0.03	-0.05	0.11	0.04	-0.34**
		(0.15)		(0.15)		(0.18)	(0.18)	(0.32)		(0.16)
Subindex: Communication with managers									-0.06	0.19
										(0.14)
Subindex: Raising suggestions	-0.10	-0.01	0.02	0.01	-0.11	-0.02	0.02	0.04		
		(0.07)		(0.08)		(0.10)	(0.06)	(0.16)		
Subindex: Suggestion box communication	-0.06	0.13	-0.03	0.01	-0.13	0.26	0.17	-0.29	-0.23	0.57***
		(0.13)		(0.11)		(0.21)	(0.20)	(0.32)		(0.19)
Subindex: Submission of suggestions	-0.06	0.12	-0.01	0.05	-0.04	0.05	-0.03	0.14	-0.05	0.04
		(0.15)		(0.14)		(0.17)	(0.22)	(0.29)		(0.18)
Number of Observations	200		200		200				97	

Notes: Table K.27 shows the ITT effects of the lunch and huddle treatment on the workers' (columns 2 to 8) and supervisors' (column 10) perceived communication practices. Individual index components are shown in Appendix Section G. Floor fixed effects and stratification variables are included in the regressions. Standard errors are clustered at the pair level in column 2, and at the line level in all other columns. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.28. HETEROGENEOUS TREATMENT EFFECTS BY WORKERS' SKILLS

	Quantity			**Quality**		
	(1)	(2)	(3)	(4)	(5)	(6)
High-skilled worker	0.022**	0.093***	0.014	0.091	0.42***	-0.0087
	(0.011)	(0.025)	(0.014)	(0.073)	(0.15)	(0.10)
Lunch	0.48***		0.39***	2.40***		1.95***
	(0.018)		(0.029)	(0.11)		(0.20)
Lunch * High-skilled worker	0.040		0.087**	0.17		0.47*
	(0.025)		(0.040)	(0.16)		(0.26)
Huddle		0.014	-0.013		0.023	-0.086
		(0.023)	(0.0083)		(0.13)	(0.053)
Huddle * High-skilled worker		-0.041	0.0054		-0.18	0.051
		(0.031)	(0.013)		(0.19)	(0.11)
Lunch * Huddle			0.038			0.12
			(0.036)			(0.24)
Lunch * Huddle * High-skilled worker			-0.075			-0.35
			(0.048)			(0.32)
Observations	4680	3634	3634	4680	3634	3634
R^2	0.39	0.08	0.33	0.28	0.06	0.24
Control Mean	0.01	0.21	0.01	0.04	1.06	0.08
Round FE	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES
Lunch=Lunch*High-skilled worker	0.00		0.00	0.00		0.00
Huddle=Huddle*High-skilled worker		0.27	0.28		0.48	0.31

Notes: Table K.28 corresponds to Table 4, but uses the number of processes a worker can perform instead of the experience index as a measure of a worker's skillset. "High-skilled worker" means the respective worker can perform an above-median number of processes, as measured in the skills matrices data. This data is not available for 26% of the workers, explaining the lower number of observations compared to Table 4. Columns 2, 3, 5, and 6 exclude all observations prior to the start date of huddles, i.e. the 14th of January 2020. A the bottom, I show p-values from comparing each treatment and treatment interaction against each other. Regressions include round and floor fixed effects. Standard errors are clustered at the pair level in columns 1 and 4, and at the line level in columns 2, 3, 5 and 6. Standard errors are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.29. HETEROGENEOUS TREATMENT EFFECTS BY WORKERS' OWN GENDER AND EXPERIENCE

	Quantity			**Quality**		
	(1)	(2)	(3)	(4)	(5)	(6)
Female worker	0.015**	0.0071	-0.033	0.16***	0.082*	-0.23
	(0.0076)	(0.0076)	(0.025)	(0.046)	(0.049)	(0.14)
High-exp worker		-0.0084			-0.066	
		(0.013)			(0.082)	
High-exp worker * Female		-0.0013			0.041	
		(0.010)			(0.068)	
Lunch	0.48***	0.40***		2.58***	2.19***	
	(0.023)	(0.038)		(0.15)	(0.25)	
Lunch * Female worker	-0.043*	-0.023		-0.50***	-0.46*	
	(0.026)	(0.041)		(0.16)	(0.26)	
Lunch * High-exp worker		0.12***			0.60**	
		(0.046)			(0.30)	
Lunch * Female worker * High-exp worker		0.0077			0.16	
		(0.052)			(0.33)	
Huddle			-0.0031			-0.027
			(0.031)			(0.18)
Huddle * Female worker			-0.0072			-0.039
			(0.035)			(0.20)
Observations	6352	6352	4955	6352	6352	4955
R^2	0.34	0.35	0.07	0.24	0.25	0.05
Control Mean	0.01	0.01	0.25	0.04	0.04	1.32
Round FE	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES	YES	YES
Lunch=Lunch*Female worker	0.00	0.00		0.00	0.00	
Huddle=Huddle*Female worker			0.95			0.97

Notes: Table K.29 shows heterogeneous treatment effects by worker's own gender and experience. "Female worker" is an indicator to one if the respective worker is female. Experience is measured by an Anderson (2008) index of the number of factories worked at, grade, tenure in the factory, tenure in the industry and the educational level. "High-exp worker" means that the respective worker has an above-median experience index. Regressions in columns 3 and 6 exclude observations prior to the start date of huddles, i.e. the 14th of January 2020. All regressions include round and floor fixed effects. Standard errors are clustered at the pair level in columns 1, 2, 4 and 5, and at the line level in columns 3 and 6. They are shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.30. HETEROGENEOUS TREATMENT EFFECTS BY EXPERIENCE AND FRIENDSHIP

	Quantity		**Quality**	
	(1)	(2)	(3)	(4)
Friend	0.011	-0.024	0.28*	0.13
	(0.028)	(0.060)	(0.17)	(0.33)
High-exp worker		0.19***		1.11***
		(0.049)		(0.30)
High-exp partner		0.086*		0.36
		(0.049)		(0.29)
High-exp worker * Friend		-0.0083		-0.27
		(0.078)		(0.46)
High-exp partner * Friend		0.15*		0.62
		(0.076)		(0.45)
High-exp worker * High-exp partner		-0.095		-0.44
		(0.063)		(0.39)
High-exp worker * High-exp partner * Friend		-0.083		0.13
		(0.11)		(0.64)
Observations	2140	2140	2140	2140
R^2	0.03	0.07	0.06	0.09
Control Mean	0.67	0.55	3.21	2.59
Round FE	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES
Stratification var.	NO	NO	NO	NO
High-exp worker=High-exp partner		0.01		0.01
High-exp worker=High-exp worker*Friend		0.08		0.05
High-exp partner=High-exp partner*Friend		0.58		0.69
High-exp p.*Friend=High-exp w.*High-exp p.*Friend		0.18		0.63

Notes: Table K.30 shows the effect of workers' own and their opposite seating partners' experience and friendship on the quantity and quality of suggestions. Experience is measured by an Anderson (2008) index of the number of factories worked at, grade, tenure in the factory, tenure in the industry and the educational level. "High-exp worker" means that the respective worker has an above-median experience index. "High-exp partner" means that the respective worker is seated opposite a worker with an above-median experience index. "Friend" is an indicator equal to one if the workers is seated opposite a person the worker talks to most about work-related issues (at baseline). At the bottom, I show p-values from comparing the effect of different interactions. Regressions exclude observations from workers not assigned to the lunch treatment. They include round and floor fixed effects. Standard errors are clustered at the pair level and shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.31. HETEROGENEOUS TREATMENT EFFECTS BY NEW SEATING PARTNERS

	Quantity		**Quality**	
	(1)	(2)	(3)	(4)
New partner	-0.026	0.00045	-0.027	0.22
	(0.036)	(0.043)	(0.23)	(0.25)
New partner * Previous friend		-0.14		-0.71
		(0.16)		(1.19)
Observations	2760	2760	2760	2760
R^2	0.19	0.21	0.12	0.14
Control Mean	0.67	0.63	3.26	3.01
Round FE	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES
New partner=New partner*Previous friend		0.43		0.46

Notes: Table K.31 shows heterogeneous treatment effects by workers assigned to a new seating partner, a non-friend, after two lunch sessions. "New partner" is an indicator equal to one if a workers is seated opposite a new seating partner in lunch session three. "Previous friend" is an indicator equal to one if a worker has been seated opposite a friend in lunch sessions one and two. I use a Difference-In-Difference specification to detect differences between rounds 1+2 and rounds 3(+4). The table only displays coefficients of the effects in round 3(+4). At the bottom, I show p-values from comparing the effect of a new partner in round three and the interaction with having been seated opposite a friend in previous rounds. Regressions exclude observations from workers not assigned to the lunch treatment. They include round and floor fixed effects. Standard errors are clustered at the pair level and shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.32. HETEROGENEOUS TREATMENT EFFECTS BY GENDER DISPOSITION AND WORKERS' EXPERIENCE

	Quantity			**Quality**		
	(1)	(2)	(3)	(4)	(5)	(6)
High-perc male peers	-0.018	0.028	0.057	-0.0097	-0.35	-0.35
	(0.027)	(0.053)	(0.098)	(0.18)	(0.37)	(0.67)
Female worker		-0.0061	-0.015		-0.79**	-0.94
		(0.058)	(0.099)		(0.39)	(0.66)
Female worker*High-perc male peers		-0.063	-0.12		0.40	0.085
		(0.061)	(0.11)		(0.41)	(0.71)
High-exp worker			0.074			0.15
			(0.11)			(0.75)
High-exp worker * Female			0.055			0.45
			(0.12)			(0.79)
High-exp worker * High-perc male peers			-0.050			-0.014
			(0.11)			(0.79)
High-exp worker * High-perc male peers*Female worker			0.091			0.61
			(0.13)			(0.87)
Observations	2140	2140	2140	2140	2140	2140
R^2	0.03	0.04	0.06	0.06	0.07	0.08
Control Mean	0.67	0.69	0.64	3.03	3.97	3.90
Round FE	YES	YES	YES	YES	YES	YES
Floor FE	YES	YES	YES	YES	YES	YES
Stratification var.	NO	NO	NO	NO	NO	NO

Notes: Table K.32 shows heterogeneous treatment effects by the gender disposition at lunch sessions and worker's experience. "High-perc male peers" means that workers are seated around an above-median percentage of male peers. The percentage of seating partners is measured as the number of male seating partners to the left, right, opposite, diagonally left and diagonally right divided by the total number of seating partners. "Female worker" is an indicator to one if the respective worker is female. Experience is measured by an Anderson (2008) index of the number of factories worked at, grade, tenure in the factory, tenure in the industry and the educational level. "High-exp worker" means that the respective worker has an above-median experience index. Regressions exclude observations from workers not assigned to the lunch treatment. They include round and floor fixed effects. Standard errors are clustered at the pair level and shown in parentheses. * p<0.1, ** p<0.05, *** p<0.01

TABLE K.33. HETEROGENEOUS TREATMENT EFFECTS ON JOB SATISFACTION

	Job satisfaction			
	(1)	(2)	(3)	(4)
Lunch	-0.30	0.081		
	(0.27)	(0.14)		
Lunch * Above-median submissions	0.85**			
	(0.37)			
Lunch * Above-median experience		0.25		
		(0.18)		
Huddle			0.19	0.15
			(0.14)	(0.14)
Huddle * Above-median submissions			-0.19	
			(0.19)	
Huddle * Above-median experience				-0.070
				(0.19)
Observations	195	195	195	195
R^2	0.24	0.22	0.22	0.19
Control Mean	-0.23	-0.23	-0.17	-0.17
Round FE	NO	NO	NO	NO
Floor FE	YES	YES	YES	YES
Stratification var.	YES	YES	YES	YES
p (L=L*Int)	0.06	0.57		
p (H=H*Int)			0.23	0.46

Notes: Table K.33 shows heterogeneous treatment effects on job satisfaction by the number of submitted suggestions and the experience index for both the lunch and the huddle treatment separately. All regressions include floor fixed effects and stratification variables. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01