Lewin at work:
Increasing productivity through group influence

Sherry Jueyu Wu*1
Elizabeth Levy Paluck*
Department of Psychology
Princeton University

*Princeton University, Department of Psychology, Princeton, New Jersey 08544

Corresponding author: 1Sherry Jueyu Wu

E-mail: jueyuw@princeton.edu

Under review, please do not circulate without permission of authors

Author’s note.

Datasets, analysis code, and pre-registration are posted at Open Science Framework (https://osf.io/d9fnh/). Study materials and additional analyses can be found in the supplementary material.
Abstract

A classic demonstration by Lewin and colleagues explored whether participatory (vs. hierarchical) meetings made work groups more productive (Lewin, 1947b). The demonstration fueled theories of group influence over individual behavior, but did not initiate an enduring social psychological tradition of experimenting with real-world groups. We revisit for the first time the “participatory group” paradigm, using a field experiment to test the impact of increased participation. The experiment involved 65 groups (1,752 workers) in a modern factory. Half of the groups were randomly assigned to a 20-minute participatory meeting once per week for six weeks, in which the group’s supervisor stepped aside and workers contributed ideas and personal goals in an open discussion of their work. The other half continued with status quo meetings in which supervisors spoke, workers listened, and a researcher observed. We found that participatory meetings led to a 10.6% average increase in individual treatment workers’ productivity, an increase that endured for 9 weeks after the experiment ended. Extending the scope of Lewin’s demonstration, we find that the brief participatory meetings also increased treatment workers’ feelings of empowerment such as job satisfaction and sense of control. Increases in productivity correlated with greater frequency of voicing opinions within the group, but not with informational gains or with particular types of goals. These results affirm the value of studying a long-term behavioral equilibrium that is maintained by individual, group, and institutional forces. The results contribute a richer theoretical understanding of participation in groups, and a pragmatic intervention for behavior change.

Word count: 250
Human social groups exert powerful influence over their members’ behavior (Allport, 1948; Tajfel, 1982; Burnes, 2007). According to Kurt Lewin, the founder of experimental social psychology, groups are so important that any attempt to change a person’s behavior will fail in the long run if it does not involve her group (Lewin, 1947a). The success of “group carried” behavior change (Maier, 1965) is due to the regularity with which individual behavior is constrained by group pressure to fit in. Individuals who change outside of their group context are likely to experience pressure from their meaningful social groups to shift their behavior back to its original pattern (Lewin, 1947a). More generally, Lewin theorized that patterns of behavior are maintained by an equilibrium of forces from the person, her social group, and her larger social environment. Psychologists who understand this equilibrium may be able to shift long-lasting patterns of behavior, rather than influencing one-time deviations from this pattern (Lewin, 1943).

Armed with this theory, Lewin launched the study of groups as a motivator for behavior change (Lewin, 1947a). Although his creative field studies did not pinpoint the precise mechanisms by which groups could change individual behavior, his work formed the basis of the field of group dynamics, a term that he coined (Kippenberger, 1998). At the time, Lewin’s move was revolutionary—to understand behavior by attending to group forces that were external to a person—since most of psychology focused on changing behavior by altering internal processes (Freud, 1962; Horney, 1950).

At present, social psychologists rarely study actual interacting social groups. Current laboratory paradigms, for example, ask individuals to imagine group identities or interactions (Kramer & Brewer, 1984; Williams, Cheung, & Choi, 2000; Prooijen, Bos, & Wilke, 2004). When actual groups are assembled in the laboratory, over three-fourths come together on a one-time basis, meaning they have no past or future (Moreland, Hogg, & Hains, 1994; see also Fiske,
2014). These paradigms preclude the study of psychological pressures generated by the group and the group’s larger social environment. Studies of actual interacting groups can be found in other areas of psychology, but the methods used are predominantly descriptive and rarely allow for causal inference (e.g., Hackman, 1990; 2002; for review, see Mathieu, Hollenbeck, Knippenberg, & Ilgen, 2017). These methods stand in stark contrast to the early decades of Lewinian research, in which experiments with face-to-face social groups uncovered powerful phenomena like group conformity (Asch, 1951), groupthink (Janis, 1972), minority influence (Moscovici, 1969), and group social norms (Sherif, 1936; Lewin, 1947b).

Today, Lewin’s legacy is often associated with the idea that small changes produce big effects (Cohen, 2011; Thaler & Sunstein, 2009). This is only a partial reading of his body of work. His field interventions were designed to affect one particular force in the equilibrium of individual, group, and societal forces (such as a person who served as a “gatekeeper,” a social group dynamic, or an institutional rule), so that he could observe whether a change in that force could shift the entire equilibrium (Lewin, 1947b). An intervention to affect that initial force, such as Lewin’s group dynamics intervention, was often designed to be strong, not subtle. A modern psychologist might describe Lewin’s interventions as “compound,” meaning that they include not one but two or three steps, or that they activate more than one psychological process at a time. To the modern psychologist, these interventions may not seem “small,” but they are small within Lewin’s perspective on the general equilibrium of behavior.

The present work follows that tradition, and seeks to revive an interest in experimentation that is aimed at testing theories about the full ecology of forces driving individual behavior in the world. We specifically study the paradigm of participatory group meetings, in which members of a work group are invited to discuss work challenges with one another and develop individual
work goals. In their famous Harwood pajama factory demonstration (Marrow, 1969), Lewin and colleagues reported a substantial positive effect of these group-based participatory meetings on worker productivity. This work is now part of the canon of studies that changed the course of social psychology. However, a close examination shows that its lessons about participation may not be the most reliable; the work was conducted with a single factory work group that was selected for its high performance, and the data were analyzed without modern statistical tools.

By revisiting this study in a multinational apparel manufacturing factory in China, we aim to provide a statistically and methodologically powerful test of Lewin’s foundational idea that modifying the participatory dynamics of groups can change and then sustain new behavioral patterns for their individual members. We also expand the scope of this idea by exploring attitudinal effects, possible mechanisms of change from participation to behavior, and by testing how long the effect, if any, can last. Our field experiment introduces participatory meetings to a random subset of 65 work groups over the course of 6 weeks. In total, the intervention involves 1,752 factory workers, with repeated behavioral and attitudinal measurement over the course of approximately 9 months. After testing for the behavioral effects of this intervention, we use quantitative and qualitative data to explore hypotheses about processes that could lead to behavior change: gaining information from the group, goal-setting, or having one’s voice heard by the group.

Lewinian Theory of Group-based Behavioral Change

In his group dynamics work, Lewin sought to formalize a model that harnessed the powerful forces of a social group to change and sustain individual behavioral patterns. He developed a 3-step model of individual behavior change, which described *unfreezing*, *moving*, and *refreezing* a behavioral pattern. For example, to decrease an individual’s racially biased
behavioral patterns, Lewin theorized it was first necessary to unfreeze the pattern through an emotionally stirring or unusual group-based event (Lewin, 1947b). He used group interventions called “T-groups,” or training groups, in which emotional topics such as prejudice and discrimination were discussed with a moderator. Moving the behavioral pattern toward a new equilibrium was the next step, and Lewin theorized that the unfreezing event could itself create an opportunity for the individual to learn and change. For example, individuals might learn from other people in their T-group discussions, and desire to be more like them.

Lewinian theory posits that social groups are necessary to refreeze a person’s new behavioral pattern. Individual actors might be able to unfreeze and move a person’s behavioral patterns, but only social groups can sustain that change by reinforcing the new pattern (Lewin, 1947a, p. 36). For example, groups might adopt a new behavior alongside an individual or interpret the new behavior as consistent with their ongoing group identity (Lewin, 1947a, p. 199; Miller & Prentice, 2016). In addition to the T-group, Lewin developed another type of group intervention, which we term “participatory meetings1,” a type of group-based exercise used to unfreeze and move group members’ behavior (Lewin, 1947b).

Participatory meetings: The Harwood studies. The set of classic studies that initiated social psychology’s interest in the group as a motivator of behavior change were the Harwood pajama factory studies, conducted by Kurt Lewin and Alex Bavelas during the 1940’s. Lewin was invited by his PhD student Alfred J. Marrow to visit the Harwood Manufacturing Corporation’s new factory in Marion, a rural community in Virginia (Marrow, 1972; Burnes,
Lewin at Work, Wu & Paluck

2007). The plant employed roughly 350 workers, who were mostly female, inexperienced, and uneducated, and whose productivity and morale were low.

Lewin and colleagues theorized that changing worker productivity would require the motivation of a face-to-face work group to “unfreeze” the current behavioral pattern, motivate a new pattern of behavior that led to higher productivity, and “refreeze” the new pattern. They designed a participatory meeting paradigm to test with factory sewing workers (Lewin, 1947b; Maier, 1965; Marrow, 1969, 1972; Burnes, 2007). In their study, one group of workers met with Lewin’s colleague Bavelas at the start of the workday. Bavelas encouraged the group to discuss the day’s work, barriers to their productivity, and possible solutions to these barriers. Finally, he encouraged each worker in the group to announce her own individual production goals in front of the group. Lewin described these meetings, which were held three times during an experimental period, as “democratic” (Maier, 1965). They stood in contrast to standard operating procedure at the factory, by which workers were subject to a more autocratic management style and followed the rules and decisions of the management (Bavelas, 1948; Lewin 1947a).

While Lewin did not spell out the three steps of change for this intervention, one imagines that the new grouping of workers and the shift from an autocratic to a democratic management style galvanized the workers and unfroze their existing behavioral patterns. Voicing one’s opinion about work, or learning new information from conversations with colleagues, may catalyze the motivation to change behavior. Finally, a regular implementation of these meetings or aspects of these meetings, like setting goals in front of coworkers, may help to “refreeze” the new and more productive behavioral patterns.
Their results were striking: workers in the participatory meetings increased their production sharply, compared to all other workers. This reported increase in behavioral productivity held steady over time, despite the fact that the treatment group’s performance was already above the factory average (see Figure 1).

Figure 1. The Demonstration Effect of Participatory Group Meetings on Sewing Machine Operators’ Productivity (Lewin, 1947b). The study description was unclear about the duration and the end point of these participatory meetings. (Figure reproduced from Maier, 1965.)

These remarkable findings notwithstanding, several aspects of the study’s setting and methods call into question whether its effects are causal or reliable. First of all, the participatory meetings treatment and researcher observation was applied to a single group of sewing machine operators, who were hand-picked by the factory as outstanding performers. In addition, features of the Harwood factory context may have been critical to the success of the treatment. For example, other experiments conducted on the same workers during that period may have been partially responsible for the effect (including stereotype reduction and sensitivity training reported by Lewin and colleagues elsewhere; Marrow, 1969; Patnoe, 1988). Furthermore, the head of the factory, Alfred Marrow, was Lewin’s former student and a supporter of scientifically...
grounded management, enthusiastic to transform the dynamics of the hierarchy from top to bottom (Marrow, 1969). In sum, the Harwood study of participatory meetings was a highly influential study of group dynamics and behavior. But upon close examination, it seems better understood as an inspiration to social science than as reliable evidence.

To our knowledge, the participatory meeting treatments have never been replicated or extended by other researchers, although they have been adopted by corporations who invited Lewin as a consultant following World War II (Marrow, 1969). However, there has been active empirical research into the broad and related topics of teams, participation, and hierarchy since Lewin’s work. Despite this large literature, which we review below, the effects of participation in group life on productivity are far from established. Experimental studies with artificial work groups have provided mixed results, some of which contradict the optimistic findings from the Harwood studies. Actual work groups in the field are almost universally studied with correlational designs that cannot rule out alternative influences on worker behavior. The question of whether group participation drives productivity is still surprisingly open.

### Teams, Participation, and Hierarchy

A full review of the enormous literature on teams, participation, and hierarchy is beyond the scope of this paper. Below, we summarize what evidence each of these research domains supplies in response to Lewin’s original hypothesis that increased participation in work groups can lead to greater productivity.

**Teams in organizations.** While Lewin is regarded as the founder of small group research, research on group and team\(^2\) dynamics has moved from social psychology to related

---

\(^2\) Some scholars draw fine distinctions between small groups and teams, or partition groups into teams, task forces, and crews (e.g. McGrath & Gruenfeld, 1993). While these taxonomies can be useful, the distinctions between small
fields like industrial psychology, organizational behavior, and communications. There, team research has prioritized large-scale description or detailed case studies over causal inference (e.g., Allmendinger & Hackman, 1996; MacDuffie, 1995; for reviews, see Ilgen, 1999; McGrath, Arrow, & Berdahl, 2000). Specifically, the evidentiary base consists primarily of surveys and case studies\(^3\) of real world work teams, which preclude making causal inferences about the behaviors or attitudes that are examined.

A predominant goal of this research on teams is to develop normative models of productive teams (e.g. Hackman, 1990; 2002; see Kozlowski & Ilgen, 2006; Guzzo & Dickson, 1996 for reviews). Paralleling Lewin’s focus on participatory, democratic groups, two active areas of research focus on autonomous work groups (AWGs, also known as semiautonomous, self-managing, self-directing, empowered teams) and quality circles. In AWGs, team members have control over when and how the work is done (Parker, Morgeson, & Johns, 2017). Relatedly, in “quality circles,” team members are consulted for advice without having actual decision power (Cohen & Bailey, 1997). Both AWGs and quality circles highlight the role of autonomy and participation in job design, and are theorized to enhance team performance (Parker et al., 2017). However, systematic reviews reveal inconsistencies in outcomes for each type of team (Cohen & Bailey, 1997; Pasmore, Francis, Haldeman, & Shani, 1982; Parker, 2014).

A subset of research on teams focuses on team leadership, and categorizes leadership styles into two clusters: one that focuses on the tasks and goals (such as transactional, authoritative, and directive leadership), and the other that focuses on developing team members, groups and teams are blurred in much of the empirical literature (Ilgen, 1999). In the current paper, we do not differentiate between groups and teams.

\(^3\) Sometimes the term “experiment” was used to refer loosely to change or manipulation of actual work activities rather than randomized control trials. For consistency, we refer to those studies as case studies.
(such as transformational, participative, and democratic leadership; Kozlowski, Gully, McHugh et al., 1996; Kozlowski & Ilgen, 2006). A focus on developing team members is the style most relevant to Lewinian participatory groups. However, the (primarily non-experimental) literature finds that developmentally focused leaders do equally well as task-focused leaders for encouraging team performance (Judge & Piccolo, 2004; Kozlowski & Ilgen, 2006); relatedly, a meta-analysis revealed no clear advantage of democratic or authoritative leadership style for promoting team performance (Gastil, 1994).

**Participation.** In a related literature, researchers have defined “participation” as a behavioral process in which influence or decision power is shared between hierarchical superiors and their subordinates (Wagner & Gooding, 1987a, p. 241). Defined psychologically, participation is a feeling of involvement in decision processes (Ritchie & Miles, 1970; Schuler, 1980; Miller & Monge, 1986). A weakness of the literature on participation in groups, however, is its lack of theoretical resemblance between any two studies that purport to measure or manipulate “participation.” The effects of participation, broadly defined, on group members’ behavior vary from positive to null and even some negative effects (e.g. Latham & Yukl, 1976; Schuler, 1980; Richter & Tjosvold, 1980).

According to meta-analyses, some of the inconsistent findings for the effects of participation on behavior can be attributed to methodological variations. Strong correlations between participation and behavior seem to rely on individuals’ self-reports: \( r = .39 \), while studies that measure participation or behavior with multiple methods reveal a small average correlation of \( r = .12 \) (Wagner & Gooding, 1987a; 1987b; Crampton & Wagner, 1994). Most research on the effects of participation involves experimental studies with artificial laboratory
groups or correlational studies in the field. Experimental research examining causal effects on objectively measured behavior is scarce, and often under-powered.\(^4\)

One form of participation is goal-setting in a group context, which is an explicit part of the Lewinian model of participatory groups. Research on group goal-setting largely replicates findings from research on individual goals: certain types of goals promote performance (specifically, those that contain some degree of specificity and are difficult but achievable; Atkinson, 1957; Bargh, Gollwitzer, & Oettingen, 2010), as opposed to the group nature of the goal-setting. Research shows no difference between goals that are set by the group and goals that are assigned to the group, in terms of their effects on group performance (Locke & Latham, 1990, 2002; Locke, Alavi, & Wagner, 1997).

**Hierarchy.** Participation may flatten hierarchy, by sharing influence, decision power, or more general involvement across group members. A third related literature theorizes the causal effects of hierarchy on team performance in both directions: the functionalist perspective proposes that hierarchy increases team performance because of enhanced coordination processes (e.g., Halevy, Chou, & Galinsky, 2011; Magee & Galinsky, 2008; Gruenfeld & Tiedens, 2010), while the conflict perspective proposes that hierarchy decreases performance because of increased conflict among different group members (e.g., Bloom, 1999; Greer, Van Bunderen, & Yu, 2017; Bunderson & Reagans, 2011). Reviews report mixed results on the average effect of increased hierarchy (Greer, de Jong, Schouten, & Dannals, 2018), which has invited predictions about the heterogeneity of results (Anderson & Brown, 2010; Halevy et al., 2011; Tarakci, Greer, & Groenen, 2016). Relevant to the current study, two such predictions are that hierarchies

\(^4\) Most experiments manipulating participation are not sufficiently powered for, or do not analyze properly, group-based random assignment.
might enhance performance when the group is working on tasks that are simple and routine (Anderson & Brown, 2010) or when task structure includes some level of interdependency (Havely, et al., 2011). A recent meta-analysis on 54 prior studies found largely null results from those suggested moderators (i.e. complexity, interdependence; Greer et al., 2018).\(^5\)

Like the other literatures reviewed, the empirical base of this debate about the effects of hierarchy consists mostly of correlational field studies and laboratory experiments with students (e.g. Greer & van Kleef, 2010; Tost, Gino, & Larrick, 2013). Thus, little causal evidence exists to attest to the abundant theoretical propositions about the impact of hierarchy on real world work groups.

**Summary.** Overall, the literature following the Harwood studies has surprisingly little to say about whether group participation of the kind encouraged in Lewinian participatory meetings can causally change individual and average group-level behavior. There has been enormous research interest and activity in the topics of teams, participation, and hierarchy, but many of the claims about the importance of participation in teams and groups have not been backed up by an empirical database that allow causal inference. Moreover, existing laboratory experiments do not provide evidence that takes into account the full set of individual, group, and environmental forces on behavior in which Lewin was interested. Our review indicates a need for experimental research with groups in the real world: to provide a test of causal direction, and to deepen our understanding of how behavioral patterns remain stable as a result of individual, group, and societal forces.

**Hypotheses**

\(^5\) Other moderators of hierarchy’s effects on productivity included: aspects of the team structure like membership instability and skill differentiation, hierarchy mutability, and task ambiguity (Greer et al., 2018).
The current research provides a rigorous empirical test and theoretical extension of Lewin’s original hypothesis, and addresses research debates both in and outside of social psychology on teams, participation, hierarchy, and behavioral change. Using Lewin and colleagues’ original report as a guide, we design and implement a participatory meeting paradigm. Specifically, we test the effects of participatory meetings on workers’ behavior and attitudes, and investigate the length and the processes of change. We propose the following specific hypotheses (all pre-registered at http://bit.ly/2zWf3Ga).

**Behavior.** First, what is the effect of participatory meetings on behavior? Like Lewin, we predict that participatory meetings will increase individual and group productivity, compared with groups who do not participate in participatory meetings. We treat worker productivity, retrieved from the factory’s comprehensive production data, as the primary outcome in a direct replication of Lewin’s original Harwood study.

Second, how long does the effect, if any, last? Are participatory meetings able to create a new behavioral equilibrium? Extending Lewin’s original demonstration, we measured worker productivity during the 6 weeks of experimental meetings and up to 12 weeks after their cessation. Given the study’s naturalistic setting, we can ask whether broader institutional forces in the workers’ environment respond to the treatment in a way that sustains or resists a new behavioral equilibrium. For instance, the treatment may prompt factory management to assign more difficult work to groups with increased productivity, eventually erasing the treatment effect. Alternately, workers may quit more often in lower performing groups, falsely inflating the treatment effect over time. We collect data to test both possibilities.

Third, if behavior is successfully changed, what is the process through which it changes? Informational gain and voicing one’s opinions are two potential mechanisms of participatory
meeting manipulations. In our study, we measure and relate informational gains and worker “voice” to changes in worker behavior. (While not pre-registered, we also test whether the nature of goals set in the participatory meetings relate to productivity.)

**Information.** In a workplace setting, supervisors deliver information, including goals, to workers, but this information may be amplified by discussion in a participatory meeting. Is an increase in information about the task responsible for increasing productivity? Research suggests that behavior changes may be driven by the informational gains of a goal setting process (Locke & Latham, 2002; Vroom & Yetton, 1973). Other research shows that goal setting does not have direct motivational effects, but indirectly bolsters performance by stimulating increasing informational discussions (Locke, Alavi, & Wagner, 1997; Latham, Winters, & Locke, 1994).

**Voice.** Do participatory meetings change behavior by increasing the number of times that group members’ voices are expressed and heard in the group? Even though past research has not affirmed a causal relationship between voice and productivity, related research shows that opportunities to voice one’s perspective in decision-making processes enhance individuals’ perception of fairness and justice (Lind & Tyler, 1988; Thibaut & Walker, 1978), and satisfaction with authority (Tyler, 1988; Hirschman, 1970; Adhvaryu, Molina, & Nyshadham, 2017). Control-mediated theories of procedural justice suggest having a voice is viewed as a form of control that can be used to achieve favorable outcomes (Thibaut & Walker, 1978). In participation research, voicing one’s opinions is suggested to be the most agentic and productive form of participation (Vroom & Yetton, 1973).

**Attitudes.** Lewin and colleagues did not investigate any effects of the participatory meetings on attitudes, such as whether workers felt more recognition, satisfaction, and
motivation in their workplace or beyond. We extend the scope of Lewin’s original work to predict that participatory meetings could affect a cluster of related worker attitudes and preferences that we term generally “workplace empowerment.” These include job satisfaction, sense of control, happiness and well-being (“individual empowerment”), and attitudes toward group life, friendship, and perceived care and respect from the management (“group empowerment”).

One possibility is that participatory meetings directly change attitudes. During the participatory meeting process, workers may gradually observe that “more people know my name,” “I’m part of and contribute to the performance of my work group,”6 and the like. Research shows a positive relationship between participation and workplace satisfaction (Miller & Monge, 1986; Guzzo & Dickson, 1996; Parker et al., 2017). Another possibility is that worker attitudes will change following behavior change. Consistent with research on self-perception and attribution (Bem, 1972; Ross & Nisbett, 1991; Paluck, Shafir, & Wu, 2017), attitudes such as job satisfaction and sense of control may be increased after workers observe themselves becoming more productive and making more money.

**Method**

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. We pre-registered all survey items, item groupings, and analyses at the Open Science Framework ([http://bit.ly/2zWf3Ga](http://bit.ly/2zWf3Ga); see SM section E for complete survey). For both survey waves, we do not report the results for questions regarding justice and authority, as they are discussed in another paper (Wu & Paluck, in prep).

**Experimental Context**

---

6 Quotations taken from qualitative observations of our pilot work with the participatory meetings (see Method).
Our study took place at a factory in the Chinese branch of a multinational apparel manufacturer. The factory is divided into departments, such as cotton spinning, dyeing, and sewing. We selected the seven sewing departments following our pilot study (described below) because employees in these departments work in groups. Sewing workers’ monthly gross salary is a direct reflection of their individual productivity under the piece-rate payment scheme—the more each worker produces, the more she earns (workers are predominantly female). Each individual in a group works on her own task, which is related to her coworkers’ tasks. For example, one worker may be in charge of sewing the sleeves of a hoodie while another is in charge of sewing the hood pieces. Groups who coordinate well (e.g., efficiently pass on finished pieces to the next worker) can work faster; however, coordination is not the only determinant of worker productivity. We observe heterogeneity within groups, where some workers are able to work faster and earn more than others (see Figure S1 and S2 in SM section H). Workers rarely transfer to a different group after they are hired, and each sewing group has its own supervisor who oversees group work.

As part of their normal work routine, all groups have a mandatory daily morning meeting, in which the supervisor summarizes the previous day’s work performance, recommends individual and group working strategies, and announces goals for individual workers. All workers attend these group meetings.

**Observer condition (control).** To control for work groups’ awareness of the research study, each of the control groups was conspicuously monitored by an RA for the same number of meetings as treatment groups during the experimental period (throughout the paper, we refer to this condition as the observer or control condition). The RA did not encourage any change in the status quo meeting routine. She described herself as part of the research team visiting the factory
to learn management strategies from the production floors. For the duration of the experiment, RAs silently observed as supervisors led the status quo morning meetings, which were typically 20-minute lectures on the group’s production performance and on working strategies for the near future. The supervisor announced this week’s order information and goals for each individual worker at the meeting’s end in terms of the number of pieces each worker should complete, and wrote each worker’s goal on a white board where all group members could see. In our pilot research, RAs recorded notes on these meetings, including time spent talking by the supervisor and by workers. Due to invariance in these descriptions (20-minute supervisor speech, and zero worker participation), we did not collect data on worker participation during the main experiment. RAs in the main experiment also reported zero worker participation in the control group meetings.

**Participatory meetings (treatment).** The basic structure of our participatory meeting mimicked Lewin’s original design as closely as possible (see SM section A). An RA facilitated the meeting for 20 minutes, in the presence of the group’s supervisor. The RA encouraged all members of the group to participate in a discussion about production-related issues in the supervisor’s presence. Supervisors were informed in advance that they should refrain from speaking during the discussion, and in particular that they should not interrupt the workers. During the discussion time, workers were specifically encouraged to share work experiences and production strategies for their own tasks, such as how to prepare piecework, where to put finished or unwanted pieces, or the best way to pass finished pieces to the next worker in group. RAs were trained to redirect any non-work-related conversation to production-related issues. RAs set the expectation for the meetings at the start of the first treatment meeting by saying:
“We encourage everyone to speak up. Say whatever’s on your mind about your work, such as issues yesterday or in the past week, the difficulties you have at work, or things you think will help you and others. I may ask some questions, and there are no right or wrong answers. Whatever you share will be helpful for us and for the group.”

Following this discussion, the RA announced the week’s order information, so that workers could set their individual production goal for the week. The participatory meeting ended with the RA encouraging each group member to voice her own goal for the week. Each worker received a piece of paper so that she could think of a goal and write it down, and then announce her goal to the group when it was her turn. Workers gave their papers to the RA at the end of the meeting. RAs also completed a descriptive rating sheet of each meeting directly following its completion, which we used to track possible mechanisms of change (see below).

**Pilot Study**

**Sample and Method**

One year prior to the main experiment, we conducted qualitative field observation in the factory for one month, and then conducted a pilot experiment with seven treatment groups from different departments (N = 145 workers), all selected by the factory. The pilot experiment had two goals: 1) observe how our translation of Lewinian participatory meeting structure would be experienced by different kinds of work groups at the factory, and 2) collect salary data before and after the participatory meeting trial run, to address factory management’s concerns that a change in the status quo meeting structure could hamper productivity.

---

7 In the pilot participatory meeting the goal setting was structured differently because individual goals were not relevant for each type of group involved in the pilot. Instead, workers individually responded to the RA’s question about whether in general they would commit to a goal to work harder and more precisely.
Instead of holding six weekly meetings as planned for the main field experiment, the pilot study held four total weekly participatory meetings over the course of one month. The participatory meeting procedure was similar to the method described above. Our only dependent variable for the pilot study was workers’ productivity.

**Pilot Results**

We compared productivity for workers in the seven groups that were (non-randomly) assigned to the participatory meetings to the productivity of workers in all remaining factory work groups who did not hold participatory meetings ($N = 140$ groups from 19 departments, or 2,202 workers), over the course of one month. Pre-treatment productivity between these two conditions was not significantly different, $p = 0.30$ (see SM section H).

We did not find any evidence that the participatory meetings decreased productivity (our partnering factory’s main concern). By contrast, we found a statistically significant increase in monthly worker productivity among the workers in the pilot participatory meetings condition ($\beta = 35.78$, Robust $SE = 1.39$, $CI = [18.17, 53.39]$, $p = .025$; measured by piece rate salary, see main experiment Method for detail), controlling for productivity during the month prior to the treatment. This productivity increase translates to 35.78 Yuan ($5.76), extra earned for the month, or 0.73% more earnings for the workers who were in the four participatory meetings over the course of the month. This difference is estimated to be stronger when an additional month’s pre-treatment productivity is covaried, as pre-registered (Table 1).
Table 1.

*Participation in Pilot Participatory Meetings Increased Workers’ Productivity (Yuan).*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participatory meetings</td>
<td>35.78*</td>
<td>49.08*</td>
</tr>
<tr>
<td></td>
<td>(1.39)</td>
<td>(3.27)</td>
</tr>
<tr>
<td>Baseline productivity</td>
<td>0.46*</td>
<td>0.430**</td>
</tr>
<tr>
<td>(1 month prior)</td>
<td>(0.0082)</td>
<td>(0.0027)</td>
</tr>
<tr>
<td>Baseline productivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2 months prior)</td>
<td>0.19*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0067)</td>
</tr>
<tr>
<td>Constant</td>
<td>2,344*</td>
<td>1,533**</td>
</tr>
<tr>
<td></td>
<td>(40.88)</td>
<td>(21.22)</td>
</tr>
<tr>
<td>Department fixed effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>2,225</td>
<td>2,103</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.31</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Note: Robust standard errors clustered by work group in parentheses. The dependent variable is number of Yuan earned in one month during treatment. *p < 0.05; **p < 0.01; ***p < 0.001.

**Pilot Discussion**

Based on our qualitative observations in the pilot, we concluded that the main field experiment should sample medium-sized groups (eliminating small 4-person cutting groups and over 30-person quality control groups), to best enable worker participation in discussion (see SM section B for full qualitative analyses). Our qualitative results also suggested that sewing workers were able to set individual goals at the end of the meeting, as did participants in Lewin’s original demonstration.

The positive and statistically significant increase in productivity was encouraging for the research design and for the factory. Our pilot helped us to test and fine-tune the participatory meeting paradigm, which we had replicated to the best of our ability from Lewin and his
colleagues’ scattered writings. In our main experiment, we tested our full set of hypotheses with a larger and fully randomized sample of work groups.

Main Experiment

Sample and Method

We administered the participatory meetings treatment for a total of six weeks, with 18 weeks of follow-up survey and productivity measurement. A total of 65 groups (1,752 workers, none of whom participated in the pilot) from 7 sewing departments participated in the experiment; 31 groups (863 workers) were randomly assigned to the participatory meetings condition and 34 groups (889 workers) were randomly assigned to the observer condition. Below, we exclude data from 141 workers who quit their job during the 6-week experimental period, and from 78 workers who joined groups following the experiment. Our results are insensitive to the inclusion of either and both of these categories of workers (as detailed in SM section F).

Group Randomization

We randomly assigned groups from all 7 sewing departments (65 groups; \(N=1,752\)) to either participate in a weekly morning participatory meeting (participatory meetings condition) during the time of their usual meeting, or have an observer attend the usual morning meeting (observer condition). To randomize, we used a matched randomization procedure to balance and

---

8 As requested by factory management, we commenced the experiment in one department one month prior to the rest of the sample. The content and length of the experiment in this first department remained exactly the same as the other departments, displaced in time by one month. This department is included in analysis per usual, adjusting for date. We were not concerned about spillover of treatment effects from this early-start department to other departments, given our observation of minimal contact among workers in different departments.

9 We have few concerns about spillover of treatment to control groups. The building structure and work schedule allow for little communication between sewing production floors or among work groups. Workers spend most of their time in their own group’s working area during work, and have little communication with other groups during and after work.
minimize observable differences between the groups ex ante. We applied a non-bipartite matching scheme (Beck, Lu, & Greevy, 2015), matching groups within each department on their group size, average worker productivity, normal working hours, and overtime working hours. We also took into account qualitative comments from departmental supervisors on the leadership style of each supervisor to fine-tune the group matches prior to randomization (see SM section C for matching code). When groups were paired within departments, we randomly assigned one group in each pair to the participatory meetings condition, and the other to the observer condition. Three departments had an odd number of groups; we assigned the one group that did not receive a match to the observer condition, as desired by the factory. (Results are robust to the exclusion of these three groups; see Table S4 in SM section D.)

Individuals in all groups provided informed consent during a recruitment phase one month before the experiment’s commencement (we did not observe any refusals during the recruitment phase). With the cooperation of the factory’s human resources managers, researchers made oral public announcements in the sewing departments to invite workers to a study called “worker experience in the factory.” Workers were specifically told that “researchers are not part of the factory but are coming to learn management practices and offer new technologies on work-related issues. All of you are invited to take part[…] Participation is completely voluntary.”

**Experimental Procedure**

Once per week for six weeks, during the traditional morning meeting slot, Chinese research assistants (RAs; all female) facilitated the weekly participatory meetings or served as observers. Each type of meeting lasted for roughly 20 minutes before workers began their workday. RAs were graduate students from a local university, trained by the first author to follow a detailed experimental protocol. RAs were unaware of specific research hypotheses. By
the end of the experimental period, treatment groups \((N = 31)\) had experienced six weekly participatory meetings and control groups \((N = 34)\) had experienced six weekly meetings with an outside observer.

**Behavioral Outcome Measures**

**Productivity rates from factory data.** Workers in the factory are paid by piece-rate. The factory uses advanced technology that counts each worker’s finished pieces by machine in real time, providing objective, precise, and accurate measures of worker production. Piece quality is taken into account: inspected pieces that do not meet the factory’s quality standards are not counted and are passed back to individual workers to repair.

We used two types of data to evaluate workers’ productivity, our primary dependent variable. We used gross salary, which is important from the workers’ perspective, and the market value of a worker’s production, which is important from the factory’s perspective. All data were acquired from the factory’s human resources department.

**Survey Data Collection**

One week after the experimental intervention had ended, a team of 11 RAs and the first author collected individual surveys from all 1,752 members of the 65 sewing groups in the study. We repeated this survey procedure three weeks later, a full month after the intervention ended. Since a considerable part of the factory worker population is illiterate, a survey administered in writing was not feasible. Prior in-depth cognitive interviews (with factory workers who were not involved in the experiment) guided the development of our survey instrument, which combined oral questioning in a group setting with workers checking boxes on individual answer sheets.

Due to workers’ time constraints, we conducted the survey during the 1-hour meal time and provided a free meal as incentive. On average, 4–5 groups (not segregated by treatment
assignment) gathered for a mealtime survey session. Researchers read each survey question aloud, and participants marked their responses on answer sheets (featuring places for “yes” or “no”, or numbers from 1–6 on a Likert scale of agreement that was explained in advance). This procedure obviated the need to read or write Mandarin characters.

Workers were assured that there were no right or wrong answers, that the survey was confidential, and specifically that researchers would not share individual answers with the factory management. Participants did not write down their names; researchers linked survey responses to participants’ factory data with a small code on the answer sheet, which corresponded to a coded sheet of names maintained by the researchers.

Researchers described the survey as “part of a research project that investigates worker experiences in the factory.” Participants were encouraged to interrupt the researchers for question clarification, but were not allowed to look at each other’s answer sheets or discuss their answers during the session. To further ensure confidentiality, participants put their completed answer sheets in a sealed envelope and put their envelope in a box with all other surveys. As agreed in advance of the study, no identifiable survey data were shared with the factory.

**Wave 1 post-intervention survey: One week later.** The “Wave 1” survey, one week after the end of the experiment, measured multiple work-related attitudes and preferences (survey completion rate = 83.79%; 93.78% female; see Table S5 in SM section E). The survey consisted of five parts: individual job-related attitudes (**individual empowerment**), individuals’ attitudes and feelings toward the groups (**group empowerment**), mechanism measurement (**information gain and voice**), demographic information, and the manipulation check. Both exploratory and confirmatory factor analyses supported these pre-registered item groupings.
**Attitudes: Individual empowerment.** Eleven questions elicited different aspects of workers’ individual feelings of empowerment at work. We created separate pre-registered indices of specific topics, including job satisfaction (e.g., “all in all I am satisfied with my job,” and “being frustrated comes with this job,” reverse-coded), perceived sense of control at work (e.g., “there is really no way I can solve all the problems I have at work,” reverse-coded), happiness and well-being at work (e.g., “did you experience happiness during a lot of the day yesterday?”), and sense of individuation (measured by one item: “most of the people in my group know my name”).

**Attitudes: Group empowerment.** Thirteen questions investigated aspects of workers’ sense of empowerment as a group. We created pre-registered indices of specific topics, including attitudes toward their work groups (e.g., “I feel I am really part of my group,” “I have confidence and trust in my coworkers”), friendships at the factory (e.g., “I feel lonely in this factory”), and perceptions of their work group’s importance at the factory (e.g., “The factory cares about and respects us”).

**Mechanism: Information gain.** We pre-registered two possible mechanisms of change: informational gain and voice in the work group. Two survey questions elicited workers’ informational gain: “how many different gestures or strategies are you aware of that you can use to do your task?”, and “do you know who to ask for if your machine needs fixing during the order switch?” For a group-level assessment, we used research assistant estimates of the average percentage of time each treatment group spent during their participatory meetings on concrete problem solving, and a rating of how many people were involved in that discussion (1 = almost no one was engaged, 4 = almost everyone was engaged).
Mechanism: Voice. To measure voice in the work groups, we used RA ratings of how much time the group spent discussing non-informational or strategic issues, specifically: raising awareness of existing problems (without problem solving), and non-production-related issues such as food in cafeteria. For each topic, research assistants also rated how many people were involved in these discussions (1–4, same scale). In addition, we recorded two negative measures of voice: the frequency of supervisors interrupting worker discussion, and the frequency of supervisors scolding the workers (expressing negativity), and one positive measure (encouraging her or his workers to speak out) on a 1–3 scale (“never,” “a few times,” “more than several times”).

Goal-setting. One hypothesis that we did not pre-register is the idea that the content or specificity of the goals set by workers in their meetings might influence productivity. We can measure the content of the treatment workers’ goals because they wrote down their personal goals on a paper before announcing them to the group; RAs collected these papers so that team leaders would not take them and use them as reminders outside of the context of the participatory meeting. Because workers frequently did not write their names on the papers, these goals are more reliably recorded at the group and not the individual level.

Demographics. Toward the end of the survey, we measured standard demographics such as age, gender, marital status, and rural or urban origin. We also asked workers to self-report their productivity.

Manipulation check. We pre-registered two survey questions as experimental manipulation checks. The first asked to what extent workers were discussing their job with coworkers; the treatment condition should have increased this discussion a significant amount.
Second, we asked: “do you know some information about the order you are working on (e.g., order amount, deadline)?”

**Wave 2 post-intervention survey: Four weeks later.** To examine whether any changes endured beyond the end of the experimental period, we repeated a shortened version of Wave 1 four weeks after the end of the intervention with all participants, using the same procedure (we added a set of new measures about authority and justice, discussed in a separate working paper). We selected fifteen questions in total from the attitudinal constructs in Wave 1 to assess whether any attitudinal effect would sustain over a longer period of time (survey completion rate = 84.07%; 93.49% female; see SM section E).

**Analysis Strategy**

We tested the effects of participatory meetings, on worker productivity during the intervention and up to 12 weeks after, and survey responses one week and four weeks following the intervention, using linear fixed effects regressions.

Linear regressions used fixed effects for the seven departments in which the 65 groups were nested, a dummy variable indicating treatment, and a vector of pre-treatment individual covariates to improve efficiency (including pre-treatment productivity, work experience, and education). Robust standard errors clustered by group accounted for residual covariance on the group level. Thus, to estimate the average productivity for an individual worker $i$ of group $j$,

$$P_{ij} = \beta_0 + \beta_1 D_{ij} + \gamma_1 Z_{ij} + \gamma_2 H_{ij} + g_j + \mu_{ij}. \quad (1)$$

The regression coefficient $\beta_1$ represents the average causal effect of the treatment on worker productivity, as measured by $P_{ij}$ (averaged over the first 6 weeks following the start of the intervention). $D_{ij}$ refers to a binary variable of experimental manipulation randomly assigned to the participants, in which $D_{ij} = 1$ refers to the participatory meeting condition and $D_{ij} = 0$ refers
to the control condition. $Z_{ij}$ is a vector of individual-level worker characteristics that are unaffected by the treatment such as work experience and education. $H_{ij}$ denotes a vector of controls for pre-treatment productivity, broken up into 6-week averages. $g_i$ denotes a departmental fixed-effect, and $\mu$ is a zero-mean error term, assumed to be mutually independent across (but not within) groups. As two additional robustness checks, we also estimate productivity outcomes using group averages ($N = 65$) and group sums ($N = 65$), detailed in SM section D (results are consistent).

**Results**

**Balance Test**

We used a logistic regression with pre-treatment characteristics to predict treatment assignment. These characteristics included both worker demographics and work group characteristics (baseline productivity, gender composition of the group, worker experience, age, and education). The balance test revealed no significant observed differences on average between groups in the participatory meetings and observer condition (see SM section C).

**Manipulation Check and RA Meeting Descriptions**

The first author was on site to observe that all participatory meetings took place and that observers attended regular meetings for control work groups. Further, we used the reports that RAs filled out after each participatory meeting to verify that procedures had been carried out as expected and to monitor for any adverse events.

Two manipulation checks in the first survey wave each supported the hypothesis that the participatory meetings were carried out as planned. Workers in the participatory meetings condition reported more frequent discussion with group members about how to do their job well ($M_{PM} = 2.60, SD = 0.25; M_O = 2.39, SD = 0.26; \beta = 0.24, CI = [0.16, 0.33], SE = 0.04, p < .001$;
using a scale from 1 = never to 3 = a few times per week). Also, workers assigned to participatory meetings were more likely to report knowing order amount and deadlines for the production order they were working on, which were announced in the participatory meetings so that workers could form individual goals (66.42% for treatment and 56.01% for control; $\beta = 0.49$, $CI = [0.09, 0.90]$, $SE = 0.21$, $p = 0.017$; using a binary scale 1= yes and 0 = no).

From RAs’ weekly reports on the participatory meetings, we found that workers in these meetings spent on average 43.5% of discussion time on concrete problem-solving ($SD = 22.97%$), 22.0% of time on raising awareness of existing problems (without problem solving; $SD = 13.66%$), and 7.0% on discussing non-production-related issues such as food at the cafeteria ($SD = 11.80%$). Factory supervisors who were present at all meetings on average rarely interrupted the discussion ($M = 1.65$, $SD = 0.67$), reprimanded workers during discussion ($M = 1.32$, $SD = 0.54$), or encouraged workers to speak out ($M = 1.22$, $SD = 0.49$; for ratings of 1 = never, 2 = once or twice, 3 = many times).

**Worker Productivity**

During the six weeks of the treatment, workers in the participatory meetings condition were significantly more productive than workers in the control observer condition. This difference is large in monetary terms, and is robust to different measures of productivity (Table 2). Treatment workers earned on average 592.30 Yuan ($87.74$) more than the control workers over the course of six weeks ($CI = [142.50, 1042.10]$, $SE = 229.31$, $p = 0.010$). The participatory meeting effect represents a 10.63% increase from the prior six weeks in workers’ average gross salary, relative to control workers. In addition, treatment workers produced 368.76 Yuan ($54.63$) more goods in market value (measured by raw amount produced) than control workers ($CI = [56.35, 681.16]$, $SE = 159.26$, $p = .021$). This difference, which is of primary interest to the
factory, represents an 8.68% increase in average raw amount produced. Thus, participatory meetings increased productivity from both the workers’ and factory management’s perspective.

**Long-term Worker Productivity**

The productivity gains among workers in the participatory meetings condition relative to observer condition endured for 6 weeks after the experiment, a time in which none of the groups experienced a participatory meeting. For the groups previously treated with the participatory meetings, we observed sustained effects on both gross salary and raw amount produced. Workers in the participatory meetings condition earned 532.72 Yuan ($85.20) more than workers in the observer condition ($CI = [180.22, 885.22], SE = 179.70, p = 0.003), a 10.74% increase in average gross salary relative to control. Furthermore, treatment workers produced 351.28 Yuan ($56.97) more goods in market value than control workers ($CI = [79.92, 622.65], SE = 138.34, p = 0.011), a 9.41% increase in treatment workers’ market production relative to control. Again, results are robust to using departmental fixed effects and baseline covariates. The statistically significant treatment difference endured until 9 weeks following the intervention (Figure 2).
| Table 2. Productivity Change During the Six-Week Experiment Period and Sustained Productivity Change After the Experiment. |
|---|---|---|---|---|---|---|---|---|
| Dependent variable: Productivity (in Chinese Yuan) |
| | Experimental period | | | Long-term | |
| | Gross salary | Market value | Gross salary | Market value |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Participatory meetings | 584.39* | 592.30** | 362.97* | 368.76* | 491.17* | 532.72** | 320.94* | 351.28* |
| | (259.12) | (229.31) | (178.00) | (159.26) | (206.21) | (179.70) | (159.10) | (138.34) |
| Work experience | 42.87 | 32.20 | 63.72** | 50.44** |
| | (30.78) | (20.72) | (23.63) | (17.83) |
| Education | 184.19 | 119.92 | -109.21 | -95.51 |
| | (118.97) | (87.89) | (180.72) | (146.69) |
| Baseline productivity (first 6-week period) | 0.44*** |
| | (0.06) |
| Baseline productivity (second 6-week period) | 0.04 | 0.06 | 0.13* | 0.14** |
| | (0.05) | (0.05) | (0.05) | (0.05) |
| Departmental fixed effects | YES | YES | YES | YES | YES | YES | YES | YES |
| Constant | 6,310.16 | 4,043.36 | 4,800.59 | 3,086.77 | 7,656.74 | 5,860.00 | 5,761.08 | 4,229.34 |
| | (466.48) | (411.36) | (345.24) | (299.23) | (279.19) | (379.50) | (205.49) | (297.80) |
| N (clusters) | 65 | 65 | 65 | 65 | 65 | 65 | 65 | 65 |
| N (individuals) | 1611 | 1490 | 1611 | 1490 | 1561 | 1440 | 1561 | 1440 |
| Control mean estimate | 6320.64 | 6455.84 | 4902.48 | 5008.78 | 5648.30 | 5603.92 | 4522.36 | 4503.07 |

Note: The comparison condition to the participatory meetings is the observer condition. Models include full-time sewing workers paid by piece-rates. Supervisors (N = 65) and staff members (N = 76) whose productivity cannot be determined by either gross salary or market value were excluded in the productivity data analysis. *p < 0.05; **p < 0.01; ***p < 0.001.
Figure 2. Worker Productivity, Comparing Groups Using Participatory Meetings and Observer Meetings, Across a 27-Week Period.

Note: Fitted lines with 95% confidence intervals chart the complete time series of worker productivity in terms of gross salary for the participatory meetings (treatment) and observer (control) work groups—specifically, the immediate effect of the participatory meetings that emerged following the first week of intervention, and the duration of the effect for 9 weeks following the cessation of participatory meetings. Each red and blue dot represents the sum gross salary for a participating work group (red for treatment and blue for control) for one day.
Figure 2 also displays well-known seasonal trends in production patterns (Levitt & List, 2011): productivity is expected to rise steadily from February to early summer, in a post-Chinese New Year production surge, and decrease from June to October, a lackluster season for apparel manufacturing. Given these trends, the figure suggests that the intervention stabilized the treatment groups’ productivity, preventing it from decreasing as otherwise would be expected from seasonal patterns of production.

Interestingly, these behavioral results were not mirrored by participants’ self-reported change in productivity. In Wave 1, participants’ self-reported productivity (“I think my productivity has increased over the past month” on a 1–6 scale) showed no differences between treatment and control workers. All workers rated their productivity as having moderately increased ($M_{PM} = 4.56, SD = 0.47; M_O = 4.45, SD = 0.40; p = 0.32$).

**Survey Results**

**Attitudes: Individual and group empowerment.** Workers assigned to the participatory meetings reported higher individual and group empowerment, as measured by several indices. For the group of indices that we pre-specified as indicating individual empowerment, treatment workers reported significantly more job satisfaction ($\alpha = .67; M_{PM} = 4.47, SD = 0.34; M_O = 4.28, SD = 0.34; \beta = 0.17, CI = [0.01, 0.33], SE = 0.08, p = .03$), and more sense of control at work ($\alpha = .32; M_{PM} = 3.97, SD = 0.32; M_O = 3.68, SD = 0.30; \beta = 0.25, CI = [0.12, 0.38], SE = 0.06, p < .001$), compared with workers in the observer condition. One exception was happiness and well-being ($\alpha = .62$), where the difference between treatment ($M_{PM} = 4.08, SD = 0.33$) and control workers ($M_O = 3.92, SD = 0.40$) was in the expected direction, but not significant ($p = 0.10$; see Table S5 in SM).
Workers in participatory meetings expressed higher aspects of group-based empowerment as well, including favorable attitudes toward their work group, affiliation with the group and trust and confidence in group members, than workers in the observer condition ($\alpha = .80; M_{PM} = 4.67, SD = 0.23; M_O = 4.54, SD = 0.31; \beta = 0.13, CI = [0.004, 0.27], SE = 0.07, p = 0.043$). Treatment workers also felt less lonely ($M_{PM} = 2.99, SD = 0.48; M_O = 3.31, SD = 0.44; \beta = -0.31, CI = [-0.51, -0.11], SE = 0.10, p = 0.003$) and reported that the factory cared about and respected them to a greater extent than workers in the observer condition ($M_{PM} = 3.60, SD = 0.41; M_O = 3.11, SD = 0.63; \beta = 0.48, CI = [0.23, 0.73], SE = 0.13, p < .001$).

**Longitudinal Attitude Change.** In survey wave 2, four weeks after the end of participatory meetings, we repeated the manipulation check question. Even though they no longer took part in participatory meetings, treatment workers reported more frequent discussion with group members about how to do their job well than did the control workers ($M_{PM} = 2.45, SD = 0.20; M_O = 2.23, SD = 0.30; \beta = 0.24, CI = [0.13, 0.35], SE = 0.06, p < .001$).

Just as they did in Wave 1, workers assigned to participatory meetings reported more positive attitudes in the indices indicating individual empowerment. Treatment workers reported more job satisfaction ($M_{PM} = 4.06, SD = 0.29; M_O = 3.85, SD = 0.27; \beta = 0.22, CI = [0.09, 0.36], SE = 0.07, p = .001$), and more sense of control at work ($M_{PM} = 3.77, SD = 0.33; M_O = 3.53, SD = 0.39; \beta = 0.27, CI = [0.11, 0.43], SE = 0.08, p = .001$) compared to control workers. Again, there was no difference in reported happiness and well-being between treatment ($M_{PM} = 3.55, SD = 0.13$) and control workers ($M_O = 3.52, SD = 0.18; p = 0.79$).

For the group of indices indicating aspects of group-based empowerment, treatment workers continued to report that the factory cared about and respected them more ($M_{PM} = 3.42, SD = 0.40$) to a greater extent than the control workers ($M_O = 3.11, SD = 0.58; \beta = 0.34, CI =
[0.13, 0.56], $SE = 0.11, p = 0.002$). Even though treatment workers expressed more favorable attitudes toward group life and felt less lonely in Wave 1, by Wave 2 there was no difference in attitudes toward their groups ($M_{PM} = 4.19, SD = 0.27; M_{O} = 4.08, SD = 0.34; p = 0.41$) or feelings of loneliness ($M_{PM} = 3.08, SD = 0.48; M_{O} = 3.25, SD = 0.41; p = 0.13$).\textsuperscript{10}

Because we tested the average treatment effects on multiple attitudinal indices, we conducted a joint significance test against the null hypothesis that the coefficients of average treatment effects from each attitudinal treatment regression are jointly nonsignificant. As predicted, we rejected this null hypothesis: we find a jointly significant difference of the average treatment effects between workers in the participatory meeting and observer condition, $F(1, 58) = 7.24, p < .001$. This joint difference was also significant without any covariates $F(1, 63) = 6.59, p < .001$.

**Summary.** The above evidence supports our hypotheses that participatory meetings would increase productivity as well as workers’ individual and group-based sense of empowerment, including job satisfaction, sense of control, perception of respect and care from factory management, and favorable attitudes toward their work groups. The effects of participatory meetings on productivity were sustained six weeks after the intervention’s end, and the effects on attitudes were sustained at least four weeks after. These results were robust to a series of sensitivity checks (see SM sections D and F).

**Relationship Between Productivity and Attitudes**

\textsuperscript{10} During Wave 2, 78 new workers participated, having arrived after the end of the participatory meetings treatment (fewer than 10\% of our sample). When their responses are added, our significant treatment results in Wave 2 are unchanged (see SM section F). This could reflect their socialization into more productive groups; however, their small number may preclude finding a real significant difference between the group of new workers and treated workers. Because we do not know how the factory decided to assign new workers to groups, we leave their responses out of the main analyses.
We also predicted that productivity and attitudes might correlate with one another, for instance if higher productivity promoted perceived workplace empowerment outcomes. Our experimental design cannot distinguish whether the intervention directly influenced both attitudes and productivity, or whether it directly influenced one (e.g., productivity), which influenced the other (e.g., attitudes). Instead, we explore the correlations between attitudes and productivity in the same survey wave and between waves. Specifically, we use productivity and change in productivity during the intervention to predict attitudes in survey wave 1 at the end of the intervention, and we use attitudes measured in survey wave 1 to predict productivity measured 6 weeks after the intervention during survey wave 2.

Of all the attitudinal constructs, only job satisfaction was positively (though not significantly) related to productivity; moreover, this result only held for measured changes in productivity and not absolute levels of productivity. Among treatment workers, the more their productivity increased during the intervention, the more job satisfaction they reported immediately afterward \((\beta = 0.28, CI = [-0.04, 0.59], SE = 0.16, p = .08)\). The pattern did not hold for control workers. Further, among treatment workers, job satisfaction at the end of the intervention significantly predicted future increases in productivity six weeks later. The more job satisfaction treatment workers expressed in survey wave 1, the more their productivity increased six weeks after that survey \((\beta = 0.05, CI = [-0.0009, 0.10], SE = 0.025, p = .05)\). We did not find a significant relationship between absolute levels of productivity or changes in productivity and individual or group-based worker empowerment.

**Process of Change**

We explored the possibility that informational gain and having a voice in the group mediated worker productivity by analyzing RAs’ weekly ratings of the participatory meeting
discussions, including how much time the group spent discussing information about the job, or how often workers spoke up in general. These ratings were used to predict group productivity over the time course of the intervention using weekly time lag regressions on each week’s rating, and to predict post-treatment group productivity using regressions that averaged RA ratings over the entire intervention period. We also used worker self-reports on informational gain from survey wave 1.

**No effect of informational gain or goal content.** In survey wave 1, work groups assigned to the participatory meetings condition did not appear to have an informational advantage over work groups assigned to the observer condition. There was no difference in the number of gestures or strategies workers reported knowing for their sewing tasks ($M_{PM} = 1.87$, $SD = 0.41$; $M_O = 1.82$, $SD = 0.34$; $p = 0.53$) and there was no difference in workers’ knowledge of whom to contact if their machines were broken ($M_{PM} = 0.59$, $SD = 0.23$; $M_O = 0.49$, $SD = 0.26$; $p = 0.42$; using a scale with 1 = yes and 0 = no). Workers’ self-rated knowledge also did not correlate with their concurrent productivity at the time of Wave 1, or predict their productivity in the six weeks following Wave 1.

We next analyzed the variability among treatment groups’ discussion of work information in their meetings, to see whether more discussion of task-related information leads to greater productivity. Specifically, we used the combined measure of RAs’ estimates of the proportion of the meeting spent on problem solving and the proportion of workers participating in problem solving to predict group productivity (results were consistent for each estimate). Problem solving was not linked to higher group productivity, using a weekly lag of ratings and productivity (gross salary) over the course of the intervention ($\beta = 14.87$, $CI = [-12.94, 42.68]$, $SE = 14.09$, $p = .29$). Problem solving averaged over the treatment period also did not predict
productivity in the six weeks following the intervention. In fact, we found that problem solving was negatively related to productivity (gross salary) six weeks later ($\beta = -655.31$, $CI = [-1359.63, 49.02]$, $SE = 339.62$, $p = .07$).

We also analyzed whether group-level variance in productivity became smaller over time, which we might expect if workers were learning from one another and strategically coordinating around their goals. We found no difference in group-level variance in productivity comparing treatment and control groups (mean $SD: PM = 1512.27$, $SD = 547.01$; $MO = 1449.79$, $SD = 385.96$; $p = 0.52$). However, individual levels of productivity across time became more stable for treatment workers (mean $SD: PM = 64.21$, $SD = 12.70$) compared with control workers ($MO = 68.65$, $SD = 15.05$; $\beta = -5.72$, $CI = [-9.83, -1.62]$, $SE = 2.09$, $p = .006$). Thus, participatory meetings made individual but not group productivity more stable.

In terms of goals, RA records show that the workers always stated goals in terms of the number of pieces they wanted to complete each day. To test whether the size of the goal, specifically the number of pieces that the worker wanted to complete, drove the current results, we correlated the size of workers’ goals with their productivity for that week. The number of finished pieces the workers set as a goal imprecisely captures goal difficulty—some workers might work on a simpler task that can be completed at a faster rate, for example. Without knowing the difficulty of a hood vs. a sleeve, we found a small but nonsignificant correlation between goal content and productivity ($\beta = 0.05$, $p = 0.07$). Furthermore, we found that treatment workers’ goals, in terms of numbers to produce did not rise significantly from week to week ($\tau = -13.40$, $p < 0.01$), using an augmented Dickey-Fuller joint probability test of the distribution of their actual goals against a non-stationary stochastic distribution of goals over time). The change
in goal content from week to week also did not predict worker productivity within the treatment condition ($\beta = 0.05, p = 0.09$).

**Voice in the participatory meetings correlates with productivity.** Our data suggest that one important mechanism of the participatory meetings treatment was the workers’ experience of speaking up in their group, or “voice.” For participatory meeting groups, voicing opinions about production or non-production issues (without problem solving) both predicted higher group-level productivity during the intervention period ($\beta_{production} = 56.74, CI = [0.42, 113.89], SE = 27.03, p = .02; \beta_{non-production} = 64.60, CI = [25.62, 103.59], SE = 19.71, p = .001$). Voicing opinions also predicted higher group-level productivity in the six weeks following the treatment ($\beta_{production} = 1294.18, CI = [373.11, 2215.24], SE = 429.44, p = .009; \beta_{non-production} = 751.31, SE = 313.02, p = .03$).

Next, we examined whether supervisors’ encouragement of workers’ voice during the meetings similarly affected productivity (RAs’ weekly responses to: “did the supervisor—” “…intervene during the discussion?”, “…scold the workers?”, and “…encourage or praise the workers when they spoke out?”). We found that supervisors’ encouragement and praise of voice positively predicted subsequent group productivity during the intervention ($\beta = 65.75, CI = [0.74, 130.76], SE = 32.93, p = .047$) and six weeks following ($\beta = 1739.60, CI = [560.61, 2918.60], SE = 568.50, p = .006$). By contrast, supervisors’ interruption of the discussion (discouragement of voice) negatively predicted group productivity during the intervention ($\beta = -67.09, CI = [-123.95, -10.23], SE = 28.78, p = .021$), although this effect did not endure following the end of the intervention. We did not find an effect of supervisors scolding the workers.

**Why Didn’t the New Behavior “Refreeze” More Permanently?**
We found that the boost to treatment workers’ productivity endured for 9 weeks after the end of the intervention. In other words, the intervention had enduring effects, but did not permanently “refreeze” the new levels of productivity. We pose three hypotheses about this lack of permanent change, two of which we were able to test with the data at hand.

**Hypothesis 1 (individual-level): The intervention was not strong enough.** Perhaps the six-week, twenty-minute meeting treatment was simply not strong enough—it was too short, too subtle, or limited to a very localized part of the workers lives—to create enduring change. We cannot test this hypothesis without varying our intervention on any of these dimensions of “intervention strength,” so we reserve this possibility for future research.

**Hypothesis 2 (group-level): The makeup of the groups changed.** Worker turnover happens on a regular basis, and so it is possible that workers who quit shifted group dynamics and thereby affected group productivity. In order to determine whether quitting affected group productivity over the long run, we first examined whether workers quit treatment and control groups at comparable rates. In the first 12 weeks after the end of the experiment, 190 (21.57%) control workers in the observer condition quit, while 124 (14.29%) treatment workers in participatory meetings quit. A logistic regression revealed that control workers were more likely to quit their jobs than treatment workers, $\beta = -0.52$, $CI = [-0.85, -0.18]$, $SE = -0.17$, $p = .002$.

Next, we tested whether there were differential effects of quitting on group productivity. We calculated the rate of quitting for each treatment and control group, and interacted the quitting rate with treatment assignment to predict average group productivity in the first two 6-week periods after the end of the experiment. We did not find any effect of quitting on group productivity. While the treatment seems to have reduced quitting, quitting within a group is not likely responsible for the fact that the treatment did not last beyond 9 weeks.
Hypothesis 3 (institutional-level): The factory adjusted to changes and allocated difficult tasks to treatment groups. At higher levels of management, the factory accepts orders from companies all over the world. Middle management then assigns these different orders to groups. In theory, random assignment of treatment should address possible systematic differences in middle management’s task assignment for treatment and control groups. However, one could predict that after the intervention began, middle management might have observed higher levels of productivity among treatment groups and consequently allocated more difficult tasks to these high-performing groups. This may explain why the treatment effect attenuated after 9 weeks—because of institutional adjustment to real changes. To test this hypothesis, we first compared the unit price (an indicator of difficulty) of each order that was assigned to the treatment and control conditions after the start of the experiment, and second, tested whether order’s unit price was related to group productivity.

We found no difference in the unit price of the assigned order piece between treatment and control groups after the start of the experiment ($M_{PM} = 2.29, SD = 1.46; M_O = 2.36, SD = 1.47; p = 0.40$). Moreover, unit price did not consistently predict group productivity ($\beta = -0.49; p = 0.59$). In other words, this particular institutional “force” shaping group behavior does not seem responsible for the attenuation of the treatment effect after 9 weeks.

Summary. The frequency of speaking up in the group, or voice, positively predicted treatment workers’ increase in productivity. We did not find evidence that any difference in information or goal setting that resulted from the meetings drove increases in productivity. The effects of participatory meetings did not endure longer than 9 weeks after the end of the experiment; group-level worker turnover and institutional-level task adjustment are unlikely contributors to the attenuation of these treatment effects.
Discussion

The present field experiment provides the first clear evidence supporting Lewin’s original idea that participatory meetings—a group process in which workers discuss their work with one another in a non-hierarchical manner—can change behavioral patterns, specifically increasing productivity. During the 6-week experimental period, treatment workers on average increased their productivity by 11%, or $88 above their previous productivity. This increase in productivity persisted for over two months after the cessation of the participatory meetings. Workers assigned to participatory meetings also reported higher workplace empowerment, including job satisfaction, sense of control, and more favorable attitudes toward their coworkers, and were less likely to quit their job. Increases in this workplace empowerment endured for at least one month after the meetings ended.

Because we conducted the experiment in the workers’ actual workplace—scheduling the participatory and observer meetings during the factory’s regular meeting times and measuring productivity in the form of workers’ real salary—our results speak to the power of these group interactions measured against the pre-existing equilibrium of forces acting on workers’ behavior. For example, our results show that discussing work with one’s coworkers for 20 minutes each week can improve average worker behaviors and attitudes across various work tasks and supervisor styles, and even when some workers leave and new group members arrive. Our results also show that this intervention is a meaningful change, despite the fact that these workers are already highly motivated to earn as much money as possible to send home to their families. The participatory meetings also positively impacted the workers despite (or perhaps because of, as we will discuss below) a lifetime spent toward the bottom of the social and workplace hierarchy in the factory and in society more generally.
Mechanisms of Change

Why did these meetings have such a strong and enduring effect on behavior and attitudes? Lewin proposed participatory meetings as a tool for “unfreezing” and “refreezing” behavior change without testing any specific mechanism. But his insights suggest participation as the key ingredient of the participatory meetings. What exactly does participation entail? We tested multiple avenues but found no evidence that the information workers gleaned from their discussions or the goals that they set\textsuperscript{11} were responsible for their change in behavior or attitudes. However, we found that voice, measured in multiple ways as the frequency with which workers spoke in front of coworkers and supervisors, was correlated with the behavioral changes we observed.

Specifically, the amount of time that treatment workers spent talking to one another (excluding informational problem-solving time), and the extent to which most workers in the group spoke, was correlated with the group’s increase in productivity. Additionally, the more that supervisors encouraged discussion and refrained from interrupting discussion, the more productive their group. All these group level analyses suggest that the more often workers voiced their perspectives, the more productive they were. To this point, it is notable that we observed a spike in productivity among treatment groups after the first participatory meeting of the intervention (see Figure 1). The first meeting’s discussion was devoted to familiarizing the workers with the meeting structure and encouraging workers to introduce themselves to one another.

\textsuperscript{11} We can compare the main experiment’s analysis of goal setting results to those of our pilot study, which used a slightly different participatory meeting format that elicited a general goal: workers were asked to respond to a question of whether they planned to “work harder this week.” The pilot observed significantly greater productivity after workers committed to this nonspecific goal. Given these two pieces of evidence, it might be possible to think of the meeting’s goal-setting activity as another form of voice, since it required each worker to speak to the group. In other words, the importance of the goal setting activity may have been that it ensured each worker’s participation in the discussion.
another—it did not involve any problem solving with work tasks. The surge in productivity following this meeting supports the perspective that it was the meetings’ facilitation of workers’ voices, rather than the informational content of their discussion, that motivated their higher productivity in the short and longer term.

**Attitude and behavior change.** Extending beyond Lewin’s predictions, we predicted and found changes in workers’ attitudes as a result of their new participatory group discussions. It is tempting to assume that the immediate and long-term increases we identified in individual and group empowerment were responsible for workers’ increased productivity. However, we find little evidence that attitude change preceded or even correlated with behavior change. Workplace empowerment and productivity increased at the same time for treatment groups, but did not correlate, with one exception: job satisfaction weakly correlated with productivity.

According to previous research on the relationship between workplace performance and workplace attitudes or “morale” (Kahn, 1960), this lack of an association is not surprising. The relationship between workplace attitudes and behavior has been controversial over decades of research (Petty, McGee, & Cavender, 1984). Despite numerous theoretical positions on its directions, strength, and moderators, the nature of the attitude–productivity relationship remains unclear. Thus, our conclusion resembles those of classic reviews on the attitude–productivity link: that “there is no systematic relationship between productivity and such morale variables” (Kahn, 1960, p. 279).

Other evidence from this study suggests that workers’ attitudes did not change through the same process that changed behavior. In general, we found workers’ actual experiences, not their reports of their experiences, predicted behavior change. For example, workers’ actual proportion of speaking time correlated with group productivity, but not workers’ self-reports of
their voice in the workplace and in their family life. Similarly, we also found no correlation between actual productivity increase and perceived productivity increase. Workers in both treatment and control groups self-reported increases in productivity. That workers’ perceptions of their own voice do not predict behavior is somewhat surprising, given a rich literature on procedural justice that connects attitudes toward authority with individuals’ perception that they have a voice in their relationship with authorities (Tyler, 2014). However, this literature offers very little evidence on the correlation between perceived voice and actual, rather than self-reported, behavior.

One possibility is that attitudes changed as a function of experimenter demand. Perhaps workers in the participatory meetings perceived that they should respond in a more empowered “Western” style for a study that they understood to be backed by a “well-known university in the United States.” We think that this is unlikely, since: 1) control participants were given the same background about the study, 2) control groups were also “observed” each week by an RA who observed and took notes, 3) treatment and control workers took their surveys in the same room, and 4) workers knew their confidentiality was protected during the survey session. Finally, it is reasonable to think that workers were more concerned about the opinion of their employer, rather than the opinion of outside researchers. In this case, it is questionable whether the workers would perceive that reporting higher levels of individual empowerment—feelings of control at work in particular—would be desirable to factory management.

In sum, the attitude changes that we track open up questions for future research to explore. It is likely that we did not capture certain individual or group-level processes that can help explain why the participatory process shifted workers’ attitudes about their workplace, their decision-making power, and their work groups. Future studies could collect more observations of
individuals discussing and working in their groups. Our findings demonstrate the promise and power of this kind of qualitative investment, for understanding more of this intervention’s mechanisms and for aiding in interpretation of quantitative results as well.

This study speaks to some of the debates and methodological gaps in the literatures on teams, participation, and hierarchy. First, the results show a clear average benefit of participation and of flattened hierarchy, in a situation in which the tasks are simple and routine, and interdependent (Havely et al., 2011; Anderson & Brown, 2010; Greer et al., 2018). This finding contradicts some of the current ideas about the role of hierarchy and participation for group productivity. Second, this study stands apart as one of the few well-powered, cluster-randomized field experiments testing causal relationships between participation and productivity. In this respect, the study provides a template for future research. Future research should continue testing the generalizability of these participatory meetings, a topic we address below.

**Future Directions**

**Contextual variations.** Our present findings suggest other important next steps, including but extending beyond future research that investigates the mechanisms of participatory group influence. Our experiment, situated in a naturalistic field context, presents a set of findings that are ultimately about change within a specific equilibrium of individual, group, and institutional forces—a particular context. Reflecting upon which aspects of this particular context might have facilitated or limited the influence of participatory groups leads to interesting future questions.

For example, we did not find that the information shared during the workers’ discussion affected worker productivity. But perhaps information sharing would matter in other workplaces, in China and elsewhere, where worker tasks are less differentiated than those in this study’s
factory. For groups in which all members are working on the same problem, or are discussing more general challenges like multi-tasking or workplace communication, information sharing might increase productivity. Information sharing might also be important for non-work groups discussing social problems like discrimination, where sharing strategies for recognizing and preventing negative stereotypes or using unbiased language could be broadly helpful for all.

By contrast, it is possible that the present study’s context exaggerated the importance of group members’ voice in the process of change. Both the institutional setting, which featured a strict hierarchical relationship between supervisors and their young female workers, and the Chinese national setting, which features discouragement of social and political dissent, minimize regular opportunities to express opinions for workers in our sample. Future research can test whether the opportunity to express oneself is as important a mechanism in other institutional and societal contexts, where group members’ voices are more routinely recognized or encouraged.

Perhaps a factory setting in China is precisely the context where a participatory meeting would present the sharpest contrast to the established equilibrium of behavior, and thus have the greatest effect. While this is possible (though new field experimental research from the United States suggests participatory meetings have a similar effect as in China; Wu & Paluck, in prep), we point out that Kurt Lewin and colleagues developed the participatory meetings in the United States to change American workers’ behavior, in a factory headed by a CEO who was open to experimental management practices. Thus, perhaps it is not the sharp uptick in worker participation that is responsible for improved productivity, but rather the installation of regular, organized opportunities to participate, since that is the common factor between these two dramatically different settings.
**Intervention variations.** The current project follows Lewin’s legacy of studying a “compound” intervention to test a hypothesis about the balance of forces driving individual behavior in a natural environment. As foreshadowed in the introduction, the intervention can be described as compound because it included multiple elements that a psychologist would theorize as a potential mechanism of change—group discussion, an altered role for the supervisor, and an altered source of goal-setting. From other perspectives, describing the participatory meetings as compound may seem strange; in the policy world, for example, these meetings at 20 minutes once per week with no material incentives involved would be deemed “light touch” and possibly too minor to create change. Interestingly, our factory collaborators believed our intervention was “too subtle to have any effect.”

Whatever the perspective, we argue that psychologists should return to the study of these relatively more compound interventions, because they are theoretically designed to test hypotheses about the interplay of individual and social forces.

Thus, we anticipate but caution against interpreting the participatory meetings manipulation as a “messy” independent variable that needs to be reduced to its constituent parts. Not very much could be stripped from the current intervention’s proceedings without changing its meaning and significance as a work meeting. For example, would a meeting without a factory supervisor seem like a meeting to the workers? The reductionist nature of the intervention, if it were broken apart into separate components, could change its meaning to participants, perhaps rendering it incomparable to the original intervention. Taking a laboratory manipulation directly into the field or transplanting an intervention from one context to the next without adequately translating so that it carries the same social and cultural significance would likely inhibit the

---

12 In field interviews with department supervisors both before the intervention and even after the intervention, we heard that the participatory meetings were not likely to change anything, since none of the information brought up by the workers would be completely novel, and the individual goals voiced out would not deviate much from the goals set by the supervisors.
processes we wish to study (Gantman et al., 2018). In sum, field interventions are often compound, but not messy. They are built to make sense to human social groups. Rather than reducing a compound intervention into separate elements, one way to further this research would be to conduct a series of experiments with compound participatory meeting interventions. Each experiment could vary one component of the meeting, or the setting and population (for an example of this technique, see Dunning et al., in press).

We hope that this research helps to reinvigorate the study of social groups in their natural environments. This approach allows for the study of how an equilibrium of forces—from the individual, the group, and from institutions and broader society—maintains long-term behavioral patterns. These types of studies have the potential to reveal behavior change mechanisms with long-term real-world effects.
References


Supplementary Materials

Table of Contents

Section A – Participatory meetings training protocol ................................................................. 61
Section B – Qualitative observations .......................................................................................... 65
Section C – Randomization ......................................................................................................... 74
Section D – Productivity results robustness checks ................................................................. 76
  Section D1 – Robustness checks with market value as DV ....................................................... 76
  Section D2 – Robustness checks with outcomes at the average group level ............................... 79
  Section D3 – Robustness checks with outcomes at the total group level ................................... 82
  Section D4 – Robustness checks with 62 groups (intact pairs) .................................................. 85
Section E – Survey materials ...................................................................................................... 88
Section F – Survey robustness check ......................................................................................... 90
  Section F1 - Survey robustness checks including new workers and part-time workers .......... 90
  Section F2 – Survey robustness checks for missing value imputation ..................................... 92
Section G – Working hour results .............................................................................................. 94
Section H – Pilot balance test and archival analysis .................................................................. 96
Section A – Participatory Meetings Training Protocol

Participatory meetings are comprised of three parts: discussion, group leader summary, and goal setting. Each part should be connected fluently and be treated as an organic entity.

Discussion leaders should start to gather people with the help of the group leaders starting 7:30am. Each meeting lasts around 20 minutes and should end before 8am. A discussion leader’s duty is to facilitate the discussion, encourage workers to speak up and actively engage in solving production related issues jointly as a group. Discussion leaders should let workers talk for most of the time during the meeting, rather than the group leader or the discussion leader herself.

While workers engage in active discussion and goal setting, discussion leaders will facilitate. For the best discussion facilitation, always use simple and everyday Mandarin that fits the workers’ communication habits. Do not use any formal or written language expressions. In your first meetings, get a general impression of the group’s dynamics such as the gender and age composition, whether people were talkative, and what people cared about the most. Use these insights to prepare for your next meetings’ facilitation.

Discussion (around 15 minutes). A pre-arranged shuttle picks up and takes everyone to the factory before 7:30am. Each research assistant has to be on the production floor where her assigned group is located. As a discussion leader (or an observer in the control condition), we can never show up late. Greet your workers politely when you arrive on the production floor. As 7:30 approaches, start to gather workers in the group with its group leader. When every worker in the group arrives, ask them to gather in a circle and greet them warmly.

During the first meetings, discussion leaders initiate a round of self-introduction, such as names and how long the workers have been working in the factory. Discussion leaders set clear
expectations during the first meeting, and briefly repeat the expectations at the start of each subsequent meeting. The protocol for discussion leaders goes:

“My name is Zhang Xiaohong, and you can call me Xiaohong. I’m a student from Soochow University, and I am helping a professor with a project on work experience. From this week on, I will come every Monday to lead a discussion with you on work related issues during your regular morning meeting time, for a period of six weeks. We will discuss the problems you have experienced in work, and the aim is for you to work better! Our meetings are easy-going. We encourage everyone to speak up! Just voice out whatever’s on your mind about your work, such as issues yesterday or in the past week, the difficulties you have at work, or things you think will help you and others. I may ask some questions, and there’s no right or wrong answers. Whatever you share will be helpful for the group and for us. I will take some notes during the discussion for research purposes, but I will not show my notes or talk with anyone who’s not in our project team, including the factory people.”

In subsequent meetings, discussion leaders repeat the expectation, “as we all know, it’s a meeting for us to share our opinions on production related issues. I’m here to discuss with you on how to work better, rather than testing you. No worry about being right or wrong. Just say whatever you think of about work and participate!”

As a warm-up for problem solving, discussion leaders can start with easy questions such as “what type of order are you all working on today?” and “what steps are each of you in charge of?” Discussion leaders prepare and facilitate two questions for the workers to discuss. The number of discussion questions is secondary to the depth of the discussion. Though the content of discussion is flexible as long as it is work-related, we do have a module of focus for each
week during the six-week intervention period. The module and suggested discussion questions are as follows:

- **Week 1:** General feedback meeting, getting to know each other
- **Week 2:** Production speed and quantity (e.g., how to work faster? What gestures and strategies are most efficient?)
- **Week 3:** Quality control (e.g., how to avoid defects? If a defect occurs, how to most efficiently coordinate for repairing? How to self-examine that finished pieces are good before going through quality control?)
- **Week 4:** Order switch (e.g. how to shorten the adaptation period when production orders change? How to deal with issues in this fast transition period?)
- **Week 5:** Group coordination (e.g. how to increase group efficiency, such as the arrangement and transition of finished pieces? How to coordinate with the person before and after you?)
- **Work 6:** Discussion topic tailored to specific groups (e.g. if a group’s major concern is its production speed, then focus more on this topic, etc.)

**Supervisor summary (0-3 minutes).** As the pre-existing 20-minute morning meetings were led by the supervisors for managerial purposes, we left this time for them in case there are other important managerial issues group members need to know that cannot be conveyed during group discussion. The first author and the discussion leaders had reminded the supervisors to keep their post-discussion summary brief.

**Goal setting (2-3 minutes).** Towards the meeting’s end, workers are encouraged to make individual goals on their daily production. Since orders are placed by customers and have tentative amounts and deadlines for production, the discussion leaders (research assistants) tell
the workers about the orders placed by customers and specifically how many pieces each order requires and how long they have before the suggested deadline. For example, if an order placed asks for 10,000 pieces within 20 days, then a worker or a group is expected to produce around 500 pieces daily if they spread production evenly across days. The calculation is simple enough to do for workers with a Chinese elementary school education. Instead of being assigned a fixed production goal daily, workers will be given all the relevant information and encouraged to come up with a daily production goal themselves. Each worker is given a small piece of paper to do simple calculation and asked to voice out their goals in front of their group members.

In the end, discussion leaders wrap up the meeting and remind them about the following week’s participatory meeting.
Section B – Qualitative Observations

About the factory.

The study took place in the Chinese branch of a multinational apparel manufacturer, which is the largest in employee size among all branches and is located in the eastern coastal area of China. Our study population, the factory workers, were mostly young women in their twenties or thirties who migrated from rural China to the city. The factory is built on the edge of the city and far from the city center, in a location that is relatively inconvenient to reach by public transportation. Around half of the workers live in adjacent factory dorms and another half commute to work on a daily basis. From field interviews prior to the experiment, we find that these workers are eager to work, but have little education (most of them have not finished high school) or training that would allow them to get a high-skilled job in the city. Hence, they enter manual work in apparel manufacturing, which is regarded as labor-intensive and low in skill requirement. Compared with its competitors, the factory pays very well and its workers, although extremely busy, mostly receive a salary in the rank of the lower-middle class in the city.

Qualitative findings from the pilot intervention,

We recount the lessons learned from the pilot intervention (labeled a to d), followed by a series of field-note episodes within these experimental meetings, which briefly sketched parts of the meeting flow and group dynamics. Rather than a unified narrative, the following entries are a

---

13 The fact that the workers are mostly young village women with little education is very similar to the personnel composition in the Harwood factory in Lewin’ time. Whereas Harwood had around 300 workers, the current factory has thousands of workers, with quite different manufacturing scale.

14 The monthly wage of a typical worker in the experimental factory ranges from 3,000 Yuan ($483.82) to more than 7,000 Yuan ($1128.92). Some reference statistics: In 2013, the minimum wage per month in Beijing was 1,400 Yuan, about 24 percent of the 5,793 Yuan average monthly wage, as calculated by the municipal bureau of statistics. In Shanghai, the 2013 minimum wage was 1,620 Yuan, or 32.2 percent of the 5,036 Yuan average wage. Those are two of the highest income cities in China. The living standard of Suzhou is very close to Shanghai.
series of unfolding actions taken from different meetings across time. We understand that observer biases are hard to avoid in qualitative data collection (Emerson, Fretz, & Shaw, 1995), but tried to be as impartial as possible about our observations and writings.

\[
(a) \text{Workers were very quiet in the first treatment meetings, and gradually talked more and more as they got used to this participatory style of meetings.}
\]

In the first meeting with the quality control group, the workers automatically formed two strictly straight lines close to each other, one in front and one back, like soldiers in an army. The researcher told them to feel more relaxed and form a circle so that everyone could face each other. In response to this, nobody moved. The same thing happened in the first meeting with the embroidery group, where workers automatically lined up, with a sizeable distance (at least 6 feet) away from the researcher. The supervisor later told us that the workers were never required to stand in lines during meetings, but they had formed such a habit for some unknown reasons. We speculate that workers have internalized the authoritarian work style of the factory, including the hierarchical arrangements, and were ready to follow any rule of an authority without question.

Workers were reluctant to speak up in the first meetings, and refrained even from saying their names, as demonstrated in the following excerpts from field notes:

[1st meeting with the embroidery group, July 1]

\[
I^{15} \text{started the meeting with self-introduction and went around the circle asking for their names. A silence. They were reluctant to talk. The girl standing in the middle of the workers facing me said “let’s start with supervisor Wang.” Wang said I had already met}
\]

\[15 \text{Taken from field notes. “I” refers to the first author.} \]
her. A short silence and I asked the person standing right next to me to start the self-introductions. The girl looked down to the ground and quietly said her name. Then one after another. I repeated each of their names, making sure I got them right.

[1st meeting with the packing group, July 7]

“What are you working on today as a group?” I asked.

The workers looked at each other and nobody spoke. I encouraged them to speak up. A woman raised her hand and said they were packing clothes. The others murmured.

However, workers adapted to the participatory style at a faster rate than we had expected. For example, at the second meeting with the embroidery group, the workers started to smile and greet the researcher when she came in, and actively discussed the June salary payment\textsuperscript{16}. During the third and fourth meetings, most group members voluntarily expressed their opinions on production-related issues such as difficulty encountered with a new order, and how to develop good gestures with a machine.

\textbf{(b) A participatory meetings intervention was the most effective when the tasks involved some level of collaboration between workers.}

One question that remained unclear in Lewin’s work was whether a participatory meetings intervention was equally applicable to every work group. As a large part of the intervention treatment focuses on eliciting information exchange among workers on production related issues, we suspected the intervention would be most relevant for the groups which require collaborative tasks between group members, or for which an individual worker’s performance is

\textsuperscript{16} The salary payment day is on the 7\textsuperscript{th} of each month.
affected by and will affect the performance of her coworkers. It was indeed what we found in our pilot intervention. In groups that require collaboration between workers, like the sewing, packing, and cutting groups, workers expressed many constructive suggestions on what they need from the workers next to them to help themselves work faster and better, and what they could do to help their coworkers work faster and better. They had never thought of or had the chance to openly discuss working strategies like those. However, the discussion about group collaboration was not very successful for the embroidery and quality control groups, whose tasks did not necessitate collaboration between workers, as illustrated in the following episodes:

[1st meeting with the embroidery group, July 1]

“Are you working on short-sleeve or long-sleeve clothes? What kind of things are you all working on as a group today?” I asked.

“Each of us works on different things.” A worker said.

“What are each of you working on today?” I asked a second question.

“We all work on different things. Different things every day. We follow the supervisor’s assignments.”

“What are some of the strategies you’ve used for this task? Do you mind sharing with others? You know as a group, we need collaboration.”

A silence.

“Collaboration is not needed. We work on different things.” A woman quietly responded.
The discussion flow was much more fluent for the groups whose workers’ jobs were interdependent. Any work problem could be easily turned into a discussion after the first meeting.

The sewing division is the factory’s largest and most labor-intensive. Workers are organized into 20-30 person groups which work on a specific order placed by companies all over the world. For example, a group may specifically work on a purple baby one-piece while another group works specifically on a blue dress during a certain time period. Each worker in a group is assigned a step in the apparel production processes and tends to repeat the same step until the whole order is completed. As a group is vertically integrated, an individual’s work performance might affect the workers after her, even though their salary earnings only depend on individual piece rates. Nevertheless, the optimal strategy is for the group to achieve its maximum productivity so that everyone can have a stable high level of output, rather than for each individual to maximize her own profit (which results in a fluctuation of individual productivity because an individual cannot produce faster when prior steps are not finished). Coordination issues become more prominent when there is a production order switch. One problem came with new cloth patches, and inefficient coordination between workers and between different divisions. Workers complained about frequent order switches because they thought their salaries would suffer. Workers expressed grudges in the discussions. As one put it: “I can’t work fast with new tasks. And I can’t work fast unless the person before me works fast.” Workers looked surprised when they heard that actually everyone shared the same problem. A solution might be as simple as help to unpack patches:

[2nd meeting with Friday’s sewing group, July 10]
A girl standing in the middle, who is in charge of the first step of the work process, said:

“It would be great if the person after me or someone else will lend me a hand to carry the materials from carts to my working desk. The materials are too heavy for me and it slows me and also the group down.”

The woman after her nodded immediately and said she never noticed that the material-moving was slowing their performance. Another worker said she hoped others would help her to do a few pieces when she could not finish all of them in time.17 For the sewing groups, any production-related issue could be developed into a discussion. Another example follows:

[2nd meeting with Thursday’s sewing group, July 9]

“We just changed new machines four days ago. It’s hard to adapt to them.”

“I’m a lot slower under the new machine and I don’t like it.”

“The technicians are not very responsive to our needs. I asked him several times to adjust my machine, but he’s slow.”

“I have that problem too!”

“Me too.”

These responses were prompted by the question, “what are some of your production problems that you’ve encountered this week?”. It turned out that the sewing departments just changed new machines the previous week. The new machines were supposed to be better and

17 Every piece that is done by a certain individual will be counted as her own production, no matter if the pieces are assigned to her or to others. Thus helping is not “free.” However, from observation and interviews, workers seldom help others unless the group leader intervenes either because they did not know others need a help or they thought helping others waste their time.
safer than the old ones. But workers did not like them. Even though the factory had organized a lecture series from a technician on using the new machines, workers still had many problems unsolved. In the meeting workers discussed the problems they had encountered with the new machines. Hearing their voice, the supervisor focused on these problems in her supervisor summary part of the meeting, such as how to avoid leaking machine oil and how to communicate problems to the technicians.

\[d\] Putting questions in context and activating social roles elicited more responses than asking about individual experiences alone. Asking a question in an utmost concrete way was the most effective in getting responses from workers.

At the first several meetings, we encountered a problem that the workers found it difficult to describe “working strategies” in detail. Some workers thought on a very abstract level and said there were no specific strategies and everything came with some working experience. Several women mentioned that people would know the right gestures with experience but were unable to describe the process further. The researcher asked what a good gesture looked like and how to develop it. Again, the workers were unable to describe it in detail. However, when we put the question in the current context of the group and asked the question in another way, workers were able to understand the question in a concrete level and start the discussion:

[4th meeting with Saturday’s sewing group, July 25]

*The group switched order from producing a summer dress to a winter baby outfit at the beginning of this week. In the discussion, everyone said something about why switching tasks was hard for them. “We cannot work fast with new tasks. It takes time.” Several workers said. I asked whether everyone slowed down during order switch and they all*
said yes, but some workers took less time to get used to the new tasks because they have
“good gesture.”

“What is a good gesture? Do you mind sharing with the rest of us?”
A silence. Some people appeared to be thinking.

“A gesture is something that comes naturally.” A worker said.

“Previously you said more experienced workers tend to develop good gestures. For those
of you who have been here long enough, how would you teach the newer workers here?
What would you do if a new worker asks you about the good gestures?”
A woman then walked close to the machine and showed how she used a mold to help sew
a squared patch onto the front of an outfit seamlessly.

Apparently, the second way of asking the same question on “good gestures” was more
effective as it created a concrete scenario for the workers to act upon, in particular when the
social role of a “more experienced worker” in relation to a “new worker” was activated.

Similarly, when we discussed why product defects occurred and how to avoid them, the workers’
first reactions were: “I should be more careful,” “It comes with working experience. With more
working experience, workers know how to avoid the defects,” as two workers said during the
discussions. But when the researcher pointed to a specific defective piece and asked how it
occurred and how to repair it, workers focused the discussions on working strategies to fix that
specific piece as well as other more general issues on product defects.

Workers were not used to thinking analytically about problems unless provided a
concrete example or scenario. It may reflect a cross-cultural difference in people’s thinking style
(Nisbett, Peng, Choi, & Norenzayan, 2011). The factory workers are used to thinking holistically
rather than analytically. Thus in the actual intervention, we always asked a question in the most
concrete way possible and activated the context and social relations surrounding the target
question to help workers engage in extended discussions.
Section C – Randomization

Table S1. Balance check. The balance test showed there was no significant differences between participatory meetings and observer condition for any pre-treatment characteristics. Omnibus $p = 0.40$.

*Note:* $^* p < 0.05; ~ ^{**} p < 0.01; ~ ^{***} p < 0.001.$

<table>
<thead>
<tr>
<th>Condition Assignment</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td>(0.219)</td>
</tr>
<tr>
<td>DeptA3</td>
<td>-0.647</td>
</tr>
<tr>
<td></td>
<td>(1.164)</td>
</tr>
<tr>
<td>DeptB3</td>
<td>-0.532</td>
</tr>
<tr>
<td></td>
<td>(1.218)</td>
</tr>
<tr>
<td>DeptC2</td>
<td>-0.501</td>
</tr>
<tr>
<td></td>
<td>(1.193)</td>
</tr>
<tr>
<td>DeptC3</td>
<td>-0.326</td>
</tr>
<tr>
<td></td>
<td>(1.180)</td>
</tr>
<tr>
<td>DeptD2</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td>(1.181)</td>
</tr>
<tr>
<td>DeptD3</td>
<td>-0.614</td>
</tr>
<tr>
<td></td>
<td>(1.243)</td>
</tr>
<tr>
<td>Work experience</td>
<td>-0.095</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
</tr>
<tr>
<td>Baseline productivity (first 6-week period)</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Baseline productivity (second 6-week period)</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.363</td>
</tr>
<tr>
<td></td>
<td>(0.304)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.852</td>
</tr>
<tr>
<td></td>
<td>(1.344)</td>
</tr>
</tbody>
</table>
Optimal non-bipartite matching codes.

library("nbpMatching") # Beck, Lu, & Greevy, 2015

# Generate example data
my_data <- data.frame(x1 = rnorm(100),
                      x2 = rnorm(100),
                      x3 = rnorm(100),
                      x4 = rnorm(100))

# Extract variables we want to match on
match_data <- my_data[, c("x1", "x3", "x4")]

# Make distance matrix (can use `gendistance` function as well)
dist_mat <- as.matrix(dist(match_data))

# Construct matches
matches <- nonbimatch(distancematrix(dist_mat))

# Each row in `extracted_matches` is a paired-match
extracted_matches <- cbind(matches$halves$Group1.Row,
                           matches$halves$Group2.Row)
Section D – Productivity Results Robustness Checks

Section D1 – Robustness checks with market value as a dependent variable.

From section “Relationship between productivity and attitudes”.

We did not find a significant relationship between objective levels of productivity and individual or group-based worker empowerment. We then calculated the percentage change in market value produced from each worker’s baseline productivity to their productivity observed at the end of the intervention (six weeks after the start of the participatory meetings) to predict attitudes. On average, treatment workers experienced an 9.28% increase in productivity, compared to -0.62% of change in productivity among control workers ($\beta = 0.09$, CI $= [0.005, 0.170]$, SE $= 0.042$, $t = 2.09$, $p = .037$). Of all the attitudinal constructs, only job satisfaction was related to a change in productivity. We found a trending evidence on the interaction effect between condition and change in productivity: among treatment workers, the greater the change in productivity, the more job satisfaction they reported ($\beta = 0.17$, CI $= [-.07, .40]$, SE $= 0.16$, $t = 1.41$, $p = .15$). The pattern did not hold for control workers. Further, among treatment workers, job satisfaction at the end of the intervention predicted future increases in productivity six weeks later. The more job satisfaction treatment workers expressed in survey wave 1, the more their productivity increased six weeks after the survey ($\beta = 0.047$, CI $= [.001, .093]$, SE $= 0.023$, $t = 2.00$, $p = .046$).

From section “Process of change”.

No difference in informational gain. We next analyze the variability among treatment groups’ discussion of information and strategic coordination in their meetings, to see whether higher discussion of this information leads to greater productivity among these groups. Specifically, we use the estimated proportion of the discussion spent on problem solving and the
number of workers who actively participated in problem solving. Used separately as two measures, and combined multiplicatively, we find the same results. We did not find significant evidence that problem solving was linked to higher group productivity using a weekly lag of ratings and productivity (market value produced) over the course of the intervention ($\beta = 13.13$, SE = 10.33, $p = .21$). We also did not find evidence that problem solving averaged over the treatment period predicted increased productivity in the six weeks following the intervention—in fact, we found that problem solving marginally predicted decreased productivity in terms of gross salary ($\beta = -421.74$, CI = [-913.30, 69.83], SE = 237.03, $p = .09$).

We also analyzed whether group-level variance in productivity decreased over time, which we might expect if workers were learning from one another and strategically coordinating around their goals. We found no difference in group-level variance in productivity (market value) comparing treatment and control groups (means of SD: $M_{PM} = 1143.09$, SD = 409.99; $M_0 = 1118.02$, SD = 291.99; n.s.). However, individual-levels of productivity across time became more stable for treatment workers (mean of SD: $M_{PM} = 28.88$, SD = 9.65) compared with control workers ($M_0 =39.56$, SD = 12.01), $\beta = -10.38$, CI = [-13.79, -6.97], SE = 1.74, $t = -5.97$, $p < .001$). Thus participatory meetings did not make groups less variable, but did make individual production more stable.

*Voice in the participatory meetings correlates with productivity.* We found that voicing one’s opinions was significantly correlated with subsequent productivity (market value), on the group level. Both voicing opinions about production issues and about non-production-related issues marginally predicted higher group-level productivity during the treatment ($\beta_{production} = 32.21$, CI = [-2.57, 67.00], SE = 17.62, $t = 1.83$, $p = .07$; $\beta_{non-production} = 21.70$, CI = [-0.51, 43.90], SE = 11.25, $t = 1.93$, $p = .06$). Voicing opinions about production and non-production
related issues also both predicted higher group-level productivity in the six weeks following the treatment ($\beta_{\text{production}} = 520.75$, CI = [112.31, 929.19], SE = 196.95, $t = 2.64$, $p = .01$; $\beta_{\text{non-production}} = 294.03$, CI = [-30.37, 618.44], SE = 156.42, $t = 1.88$, $p = .07$).

In the positive direction, we found that group leaders’ encouragement and praise of workers’ participation in the meetings significantly predicted subsequent group productivity during the intervention ($\beta = 48.13$, SE = 24.49, $t = 1.97$, $p = .05$). In the negative direction, group leaders’ interruption in the discussion (discouragement of voice) negatively predicted group performance during the intervention ($\beta = -46.36$, SE = 21.15, $t = -2.19$, $p = .03$). The positive pattern extended to productivity findings six weeks following the intervention; group leaders’ frequency of encouragement significantly predicted greater group productivity ($\beta = 1244.97$, SE = 391.26, $t = 3.18$, $p = .004$). There were no significant predictive effects of other group leader behaviors following the intervention. We did not find a relationship between the frequency of leader scolding workers during treatment meetings and group productivity.
Section D2 – Robustness checks with outcomes at the average group level, using group averages as a dependent variable.

In addition to individual treatment effects (error clustered at the group level), we estimate group-level outcomes as robustness checks, which takes the group averages of each outcome variables as one data point (i.e., there are 65 data points for each regression). The group-level robustness checks will take the form of

\[ P_t = \beta_0 + D_t \beta_1 + H_t \gamma_1 + g_t + \mu_t. \]  \hspace{2cm} (2)

where for group \( t \), \( \beta_1 \) represents the average causal effect of the treatment on group average productivity, as measured by group average productivity \( P_t \) during and after the 6-week period of the experiment. \( D_t \) refers a binary variable of experimental manipulation randomly assigned to groups (1 = treatment participatory meetings, 0 = control observer meetings), and \( H_t \) denotes a vector of baseline controls for pre-treatment group productivity, broken up into 6-week averages. \( g_t \) denotes a departmental fixed-effect. \( \mu \) is a zero-mean error term, assumed to be mutually independent across groups.

Average worker productivity.

In the six-week period of the experiment, workers who participated in weekly participatory meetings were significantly more productive than workers who had a weekly outside observer at their meeting. The difference is large in monetary terms, and is robust to different measures of productivity (refer to Table S2). Treatment workers earned 561.29 Yuan ($81.50) more than the control workers over the course of six weeks (CI = [42.95, 1079.62], \( t = 2.17, p = 0.03, SE = 258.31 \)). This equals 8.49% of workers’ average gross salary during this time period. In addition, treatment workers produced 323.71 Yuan ($47.0) more goods in market
value (measured by raw amount produced) than control workers (CI = [-37.49, 684.91], SE = 180.00, \( t = 1.80, p = .078 \)). This difference, which is of primary interest to the factory, equals 6.37% of workers’ average raw amount produced. Thus, participatory meetings increased productivity from both the workers’ and factory management’s perspective. These results are robust to using departmental fixed effects and baseline covariates.

**Long-term average worker productivity.**

The productivity gains among workers in the participatory meetings condition relative to the observer condition sustained for at least 6 weeks after the experiment, during which all workers returned to their previous regular morning meeting schedule without group discussion. The sustained gains were observed for both gross salary and raw amount produced. Workers who participated in participatory meetings earned 617.14 Yuan (89.61 USD) more than workers in the observer condition (CI = [190.02, 1044.25], SE = 212.85, \( t = 2.90, p = 0.005 \)), which equaled 10.41% of workers’ average gross salary. Furthermore, treatment workers produced 407.05 Yuan (59.10 USD) more goods in market value than control workers (CI = [81.48, 732.62], SE = 162.24, \( t = 2.51, p = 0.015 \)), which equaled 8.65% of workers’ average raw amount produced. Again, results are robust to using departmental fixed effects and baseline covariates (see table S2).
Table S2. Productivity at the average group level.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Experimental period</th>
<th>Long term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross salary (1)</td>
<td>Market value (2)</td>
</tr>
<tr>
<td>Participatory meetings</td>
<td>772.9* (296.5)</td>
<td>561.3* (258.3)</td>
</tr>
<tr>
<td>Baseline productivity (first 6-week)</td>
<td>0.3* (0.2)</td>
<td>0.2 (0.2)</td>
</tr>
<tr>
<td>Baseline productivity (second 6-week)</td>
<td>0.1 (0.1)</td>
<td>0.2 (0.1)</td>
</tr>
<tr>
<td>Constant</td>
<td>5,947.0*** (447.0)</td>
<td>4,506.1*** (772.1)</td>
</tr>
<tr>
<td>Observations</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>

Note: *p < 0.05; **p < 0.01; ***p < 0.001.
Section D3 – Robustness checks with outcomes at the total group level.

Total group productivity.

In the six-week period of the experiment, work groups that participated in weekly participatory meetings were significantly more productive than work groups that had a weekly outside observer at their meeting. The difference is large in monetary terms, and is robust to different measures of productivity (refer to Table S3). Treatment groups earned 18,612.42 Yuan ($2,808) more than the control groups over the course of six weeks (CI = [3,841.58, 33,383.26], SE = 7,360.96, \(t = 2.53, p = 0.01\)). In addition, treatment groups produced 12,551.78 Yuan ($1,894) more goods in market value (measured by raw amount produced) than control groups (CI = [1,904.19, 23,199.37], SE = 5,306.16, \(t = 2.37, p = .02\)). Thus, participatory meetings increased productivity from both the workers’ and factory management’s perspective. These results are robust to using departmental fixed effects and baseline covariates.

Long-term total group productivity.

The productivity gains among groups in the participatory meetings condition relative to observer condition sustained for at least 6 weeks after the experiment, during which all groups returned to their previous regular morning meeting schedule without group discussion. The sustained gains were observed for both gross salary and raw amount produced. Groups who participated in participatory meetings earned 18179.96 Yuan ($2874.43) more than groups in the observer condition (CI = [4002.97,32356.95], SE = 7074.19, \(t = 2.57, p = 0.013\)). This equals 11.11% of workers’ average gross salary during this time period. Furthermore, treatment groups produced 12097.43 Yuan ($1912.92) more goods in market value than control groups (CI = [1919.95, 22274.91], SE = 5078.47, \(t = 2.38, p = 0.021\)), which equaled 9.71% of work groups’
average raw amount produced. Again, results are robust to using departmental fixed effects and baseline covariates (see Table S3).
Table S3. Productivity at the total group level.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Experimental period</th>
<th>Long term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross salary</td>
<td>Market value</td>
</tr>
<tr>
<td>Participatory meetings</td>
<td>30,582.6</td>
<td>18,180.0*</td>
</tr>
<tr>
<td></td>
<td>(17,162.4)</td>
<td>(7,074.2)</td>
</tr>
<tr>
<td>Baseline productivity (first 6-week)</td>
<td>0.9***</td>
<td>0.9***</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>Baseline productivity (second 6-week)</td>
<td>-0.1</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
</tr>
<tr>
<td>Department fixed effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>120,516.7***</td>
<td>45,685.6***</td>
</tr>
<tr>
<td></td>
<td>(25,869.7)</td>
<td>(12,433.8)</td>
</tr>
<tr>
<td>Observations</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>

Note: *p < 0.05; **p < 0.01; ***p < 0.001.
Section D4 – Robustness checks with 62 groups (intact pairs).

We applied a non-bipartite matching scheme within each department and randomly assigned groups from all seven departments ($N = 65$) to either the treatment or the control condition. For departments with an odd number of groups, we assigned the one group that did not achieve a match to the observer condition, as was desired by the factory. Here we report results excluding those three unmatched departments. The results were consistent.

Worker productivity.

In the six-week period of the experiment, workers who participated in weekly participatory meetings were significantly more productive than workers who had a weekly outside observer at their meeting. The difference is large in monetary terms, and is robust to different measures of productivity (refer to Table S4). Treatment workers earned 530.22 Yuan ($80) more than the control workers over the course of six weeks (CI = [466.25, 994.18], SE = 236.52, $t = 2.24$, $p = 0.025$). This equals 9.38% of workers’ average gross salary during this time period. In addition, treatment workers produced 328.92 Yuan ($50.00) more goods in market value (measured by raw amount produced) than control workers (CI = [8.29, 649.55], SE = 163.45, $t = 2.01$, $p = .044$). This difference, which is of primary interest to the factory, equals 7.60% of workers’ average raw amount produced. Thus, participatory meetings increased productivity from both the workers’ and factory management’s perspective. These results are robust to using departmental fixed effects and baseline covariates.

Long-term worker productivity.

The productivity gains among workers in the participatory meetings condition relative to observer condition sustained for at least 6 weeks after the experiment, during which all workers returned to their previous regular morning meeting schedule without group discussion. The
sustained gains were observed for both gross salary and raw amount produced. Workers who participated in participatory meetings earned 548.73 Yuan ($83.00) more than workers in the observer condition (CI = [174.66, 923.80], SE = 191.19, \( t = 2.87, p = 0.004 \)), which equaled 10.09% of workers’ average gross salary. Furthermore, treatment workers produced 368.10 Yuan ($56.00) more goods in market value than control workers (CI = [80.62, 655.57], SE = 146.54, \( t = 2.51, p = 0.012 \)), which equaled 8.99% of workers’ average raw amount produced. Again, results are robust to using departmental fixed effects and baseline covariates (see Table S4).
Table S4. Robustness check – worker productivity with 62 groups.

<table>
<thead>
<tr>
<th></th>
<th>Experimental period</th>
<th></th>
<th></th>
<th>Long term</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gross salary (1)</td>
<td>Market value (2)</td>
<td>Gross salary (5)</td>
<td>Market value (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participatory meetings</td>
<td>540.46*</td>
<td>340.64</td>
<td>524.63*</td>
<td>356.19*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(266.86)</td>
<td>(182.35)</td>
<td>(216.57)</td>
<td>(166.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work experience</td>
<td>38.28</td>
<td>28.55</td>
<td>71.07**</td>
<td>56.37**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(32.51)</td>
<td>(24.91)</td>
<td>(18.76)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>162.32</td>
<td>104.50</td>
<td>-105.06</td>
<td>-91.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(118.30)</td>
<td>(87.51)</td>
<td>(181.17)</td>
<td>(147.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline productivity</td>
<td>0.45***</td>
<td>0.41***</td>
<td>0.32****</td>
<td>0.32***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(first 6-week)</td>
<td>(0.06)</td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline productivity</td>
<td>0.05</td>
<td>0.06</td>
<td>0.14*</td>
<td>0.15**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(second 6-week)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6,333.20***</td>
<td>4,051.23***</td>
<td>3,083.37***</td>
<td>5,870.74***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(468.87)</td>
<td>(416.00)</td>
<td>(304.98)</td>
<td>(283.91)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *p < 0.05; **p < 0.01; ***p < 0.001.
**Section E – Survey materials**

*Table S5. Survey items.*

<table>
<thead>
<tr>
<th>Category</th>
<th>Item</th>
<th>Repeated in survey wave 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job satisfaction</td>
<td>All in all I am satisfied with my job</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>In general, I don’t like my job*</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>I often think about quitting*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I am looking for a new job*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Being frustrated comes with this job*</td>
<td></td>
</tr>
<tr>
<td>Sense of control</td>
<td>To what extent do you have control over what happens on your job?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>I sometimes feel I am being pushed around in my life*</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>There is really no way I can solve all the problems I have at work*</td>
<td>Yes</td>
</tr>
<tr>
<td>Happiness and wellbeing</td>
<td>Did you experience happiness during a lot of the day yesterday?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>All things considered, how satisfied are you with your life as a whole these days?</td>
<td>Yes</td>
</tr>
<tr>
<td>Sense of individuation</td>
<td>Most of the people on my group know my name.</td>
<td></td>
</tr>
<tr>
<td>Group attitudes</td>
<td>How do you like your coworkers?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>How often do you socialize with your coworkers during work?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I feel I am really part of my group.</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>I have confidence and trust in my coworkers.</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>I like my group.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I feel that in the factory, everyone’s part of a big family.</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>I get involved to benefit my work group.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I help others in my work group learn about the work.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I assist others in my group with their work for the benefit of the group</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The work we do here is important to the factory</td>
<td></td>
</tr>
<tr>
<td>Felt care and respect</td>
<td>The factory (“the higher” as in Chinese) cares about and respects us.</td>
<td>Yes</td>
</tr>
<tr>
<td>Loneliness</td>
<td>Think about your good friend(s) in the factory, and list the number of people you can go to at the factory when you have problems because they will help you. (list #)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I feel lonely in this factory*</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: Items are translated into Chinese and back-translated to English by two independent Chinese-English bilingual speakers. Most of the items were piloted with an independent group of workers. Asterisks signify items that are reverse coded.
Table S6. Descriptive statistics for survey constructs.

<table>
<thead>
<tr>
<th></th>
<th>Wave 1</th>
<th></th>
<th>Wave 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment</td>
<td>Control</td>
<td>Treatment</td>
<td>Control</td>
</tr>
<tr>
<td>Job satisfaction</td>
<td>4.473</td>
<td>(0.340)</td>
<td>4.286</td>
<td>(0.345)</td>
</tr>
<tr>
<td></td>
<td>665</td>
<td>685</td>
<td>3.842</td>
<td>(0.309)</td>
</tr>
<tr>
<td>Individual empowerment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Happiness and well-being</td>
<td>4.075</td>
<td>(0.329)</td>
<td>3.945</td>
<td>(0.493)</td>
</tr>
<tr>
<td></td>
<td>665</td>
<td>659</td>
<td>3.836</td>
<td>(0.188)</td>
</tr>
<tr>
<td>Sense of control</td>
<td>3.966</td>
<td>(0.321)</td>
<td>3.681</td>
<td>(0.302)</td>
</tr>
<tr>
<td></td>
<td>664</td>
<td>658</td>
<td>3.528</td>
<td>(0.393)</td>
</tr>
<tr>
<td>Group attitudes</td>
<td>4.657</td>
<td>(0.222)</td>
<td>4.533</td>
<td>(0.313)</td>
</tr>
<tr>
<td></td>
<td>666</td>
<td>659</td>
<td>4.117</td>
<td>(0.531)</td>
</tr>
<tr>
<td>Group empowerment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loneliness</td>
<td>2.989</td>
<td>(0.476)</td>
<td>3.296</td>
<td>(0.445)</td>
</tr>
<tr>
<td></td>
<td>661</td>
<td>652</td>
<td>3.238</td>
<td>(0.405)</td>
</tr>
<tr>
<td>Perceived respect</td>
<td>3.601</td>
<td>(0.411)</td>
<td>3.087</td>
<td>(0.623)</td>
</tr>
<tr>
<td></td>
<td>659</td>
<td>654</td>
<td>3.088</td>
<td>(0.574)</td>
</tr>
<tr>
<td>Frequency of communication</td>
<td>2.597</td>
<td>(0.253)</td>
<td>2.388</td>
<td>(0.261)</td>
</tr>
<tr>
<td></td>
<td>663</td>
<td>660</td>
<td>2.233</td>
<td>(0.322)</td>
</tr>
</tbody>
</table>
Section F – Survey Robustness Check

Section F1 - Survey robustness checks including new workers.

Survey results.

Attitudes: Individual and group empowerment. Workers assigned to the participatory meetings reported more individual and group empowerment. For the loose group of measures designated as individual empowerment, treatment workers reported significantly more positive attitudes, including more job satisfaction (M_{PM} = 4.05, SD = 0.29; M_{O} = 3.84, SD = 0.25; β = 0.21, CI = [0.08, 0.33], SE = 0.06, p = .001), and more sense of control at work (M_{PM} = 3.77, SD = 0.32; M_{O} = 3.56, SD = 0.35; β = 0.25, CI = [0.09, 0.41], SE = 0.08, p = .002), compared with workers in the observer condition. One exception was happiness and well-being, where there was no difference between treatment (M_{PM} = 3.91, SD = 0.21) and control workers (M_{O} = 3.84, SD = 0.18, p = 0.06).

Workers in participatory meetings expressed more favorable attitudes toward their work group, including affiliation with the group and trust and confidence in group members, than workers in the observer condition (M_{PM} = 4.68, SD = 0.23; M_{O} = 4.55, SD = 0.30; β = 0.13, CI = [0.01, 0.26], SE = 0.07, p = 0.042). Treatment workers also felt less lonely (M_{PM} = 2.98, SD = 0.45; M_{O} = 3.32, SD = 0.34; β = -0.33, CI = [-0.52, -0.14], SE = 0.10, p < .001) and reported that the factory cared about and respected them to a greater extent than workers in the observer condition (M_{PM} = 3.65, SD = 0.41; M_{O} = 3.11, SD = 0.61; β = 0.53, CI = [0.29, 0.78], SE = 0.13, p < .001).

Longitudinal attitude change. In the second wave of the survey four weeks after the end of participatory meetings, we repeated the manipulation check question. Even though they no longer participated in participatory meetings, workers in the participatory meetings condition
reported more frequent discussion with group members about how to do their job well than did the control workers (M<sub>PM</sub> = 2.44, SD = 0.20; M<sub>O</sub> = 2.22, SD = 0.27; β = 0.22, CI = [0.12, 0.32], SE = 0.05, p < .001).

Just as they did in Wave 1, workers assigned to participatory meetings reported higher job satisfaction (M<sub>PM</sub> = 4.05, SD = 0.29; M<sub>O</sub> = 3.84, SD = 0.25; β = 0.21, CI = [0.08, 0.34], SE = 0.06, p = .001), and more sense of control at work (M<sub>PM</sub> = 3.77, SD = 0.32; M<sub>O</sub> = 3.56, SD = 0.35; β = 0.25, CI = [0.09, 0.41], SE = 0.08, p = .002) compared to control workers. Again, there was no difference in reported happiness and well-being between treatment (M<sub>PM</sub> = 3.91, SD = 0.21) and control workers (M<sub>O</sub> = 3.84, SD = 0.18; p = 0.11).

Treatment workers continued to feel that the factory cared about and respected them (M<sub>PM</sub> = 3.47, SD = 0.39) to a greater extent than the control workers did (M<sub>O</sub> = 3.13, SD = 0.55; β = 0.38, CI = [0.17, 0.58], SE = 0.11, p < .001). Though treatment workers expressed more favorable attitudes toward group life and felt less lonely in Wave 1, by Wave 2 there was no difference in attitudes toward their work groups (M<sub>PM</sub> = 4.19, SD = 0.28; M<sub>O</sub> = 4.10, SD = 0.32; β = 0.07, CI = [-0.07, 0.21], SE = 0.07, p = 0.31) or feelings of loneliness (M<sub>PM</sub> = 3.10, SD = 0.46; M<sub>O</sub> = 3.24, SD = 0.39; β = -0.16, CI = [-0.33, 0.02], SE = 0.09, p = .075).
Section F2 – Survey robustness checks for missing value imputation.

As a robustness check, if a worker skipped a certain question, we replace the missing response with the mean sample score of the corresponding survey item, and report the survey results with missing value imputation. The results were consistent with those from the main text.

Survey results.

Attitudes: Individual and group empowerment. Workers assigned to the participatory meetings reported more individual and group empowerment. For the loose group of measures designated as individual empowerment, treatment workers reported significantly more positive attitudes, including more job satisfaction ($M_{PM} = 4.46$, $SD = 0.29; M_o = 4.29$, $SD = 0.28; \beta = 0.16, CI = [0.03, 0.29], SE = 0.06, p = .013$), and more sense of control at work ($M_{PM} = 3.94$, $SD = 0.26; M_o = 3.70$, $SD = 0.27; \beta = 0.24, CI = [0.14, 0.34], SE = 0.05, p < .001$), compared with workers in the observer condition. One exception was happiness and well-being, where there was no difference between treatment ($M_{PM} = 4.06$, $SD = 0.27$) and control workers ($M_o = 3.93$, $SD = 0.34$, $p = 0.052$).

Workers in participatory meetings expressed more favorable attitudes toward their work group, including affiliation with the group and trust and confidence in group members, than workers in the observer condition ($M_{PM} = 4.66$, $SD = 0.18; M_o = 4.55$, $SD = 0.26; \beta = 0.11, CI = [0.01, 0.22], SE = 0.05, p = 0.029$). Treatment workers also felt less lonely ($M_{PM} = 3.02$, $SD = 0.41; M_o = 3.28$, $SD = 0.39; \beta = -0.25, CI = [-0.41, -0.09], SE = 0.08, p = .002$) and reported that the factory cared about and respected them to a greater extent than workers in the observer condition ($M_{PM} = 3.56$, $SD = 0.35; M_o = 3.15$, $SD = 0.54; \beta = 0.40, CI = [0.21, 0.59], SE = 0.10, p < .001$).
Longitudinal attitude change. In the second wave of the survey four weeks after the end of participatory meetings, we repeated the manipulation check question. Even though they no longer participated in participatory meetings, workers in the participatory meetings condition reported more frequent discussion with group members about how to do their job well than did the control workers ($M_{PM} = 2.44$, $SD = 0.17$; $M_O = 2.25$, $SD = 0.24$; $\beta = 0.18$, CI = [0.09, 0.27], $SE = 0.04$, $p < .001$).

Just as they did in Wave 1, workers assigned to participatory meetings reported higher job satisfaction ($M_{PM} = 4.05$, $SD = 0.26$; $M_O = 3.87$, $SD = 0.22$; $\beta = 0.18$, CI = [0.08, 0.27], $SE = 0.05$, $p < .001$), and more sense of control at work ($M_{PM} = 3.76$, $SD = 0.27$; $M_O = 3.55$, $SD = 0.34$; $\beta = 0.22$, CI = [0.09, 0.41], $SE = 0.06$, $p < .001$) compared to control workers. Again, there was no difference in reported happiness and well-being between treatment ($M_{PM} = 3.90$, $SD = 0.18$) and control workers ($M_O = 3.85$, $SD = 0.18$; $p = 0.15$).

Treatment workers continued to feel that the factory cared about and respected them ($M_{PM} = 3.40$, $SD = 0.36$) to a greater extent than the control workers ($M_O = 3.14$, $SD = 0.51$; $\beta = 0.26$, CI = [0.09, 0.43], $SE = 0.09$, $p = .002$). Even though treatment workers expressed more favorable attitudes toward group life and felt less lonely in Wave 1, by Wave 2 there was no difference in attitudes toward their groups ($M_{PM} = 4.18$, $SD = 0.24$; $M_O = 4.09$, $SD = 0.28$; $\beta = 0.09$, CI = [-0.02, 0.21], $SE = 0.06$, $p = 0.11$) or feelings of loneliness ($M_{PM} = 3.09$, $SD = 0.42$; $M_O = 3.23$, $SD = 0.29$; $\beta = -0.13$, CI = [-0.28, 0.01], $SE = 0.07$, $p = .074$).
Section G – Working Hour Results

A workday for the factory workers starts at 8am and ends at 5pm, with a 1-hour lunch break and 45-minute dinner break. Workers are allowed to work overtime for at most 3 hours during workdays (except Saturday). Workers are compensated for their overtime hours (1.5 times their normal working piece rate). Thus there are two types of working hours – normal working hours (8am – 5pm) and overtime hours (after 5pm and anytime on weekends).

There was no difference of normal working hours between treatment workers and control workers both during the intervention ($M_{PM} = 219.03$, $SD = 23.14$; $M_O = 221.77$, $SD = 19.33$) and 6 weeks after the intervention ($M_{PM} = 225.10$, $SD = 40.00$; $M_P = 224.90$, $SD = 39.71$). However, the treatment workers spent more total hours ($M_{PM} = 374.42$, $SD = 47.27$) than control workers ($M_O = 364.77$, $SD = 38.47$), $\beta = 11.02$, CI = [2.20, 19.84], SE = 4.50, $t = 2.45$, $p = .01$. Spreading throughout the 42 days\(^{18}\) in the 6-week period, a worker who participated in participatory meetings on average spent 15 minutes more at work than control workers. It can be interpreted as treatment workers hanging around the production floors slightly later than the control workers. The difference in total working hours persisted 6 weeks after the end of the experiment ($M_{PM} = 353.71$, $SD = 65.56$; $M_O = 340.07$, $SD = 61.21$; $\beta = 14.84$, CI = [5.34, 24.34], SE = 4.84, $t = 3.06$, $p = .002$). In other words, treatment workers on average stayed on the production floors (either at work or socialize with coworkers) 20 minutes longer than the control workers (see Table S7). No difference in working time was observed in the second 6-week period after the end of the experiment.

---

\(^{18}\) Workers usually work from Monday to Saturday, with Sunday off.
Table S7. Results for working hours.

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>During intervention</th>
<th>Long term</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Total</td>
</tr>
<tr>
<td>Treatment</td>
<td>-1.749</td>
<td>11.020*</td>
</tr>
<tr>
<td></td>
<td>(0.912)</td>
<td>(4.499)</td>
</tr>
<tr>
<td>Work experience</td>
<td>0.345</td>
<td>0.591</td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td>(0.784)</td>
</tr>
<tr>
<td>Education</td>
<td>0.224</td>
<td>4.387</td>
</tr>
<tr>
<td></td>
<td>(1.751)</td>
<td>(3.559)</td>
</tr>
<tr>
<td>Baseline work hours (first 6-week)</td>
<td>0.067***</td>
<td>0.290***</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Baseline work hours (second 6-week)</td>
<td>-0.018*</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Department fixed effects</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Constant</td>
<td>184.532***</td>
<td>247.709***</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; *** p < 0.001.
Section H – Pilot Balance Test and Archival Analysis

Table S8. Pilot balance test. The balance test showed there was no significant difference between workers assigned to the treatment participatory meetings and all the other factory workers for pre-treatment productivity in terms of salary.

<table>
<thead>
<tr>
<th>Condition assignment</th>
<th>Coefficient</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline productivity (1 month prior)</td>
<td>-0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Baseline productivity (2 months prior)</td>
<td>0.00001</td>
<td>0.0001</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.743</td>
<td>0.756</td>
</tr>
</tbody>
</table>

Note: *p < 0.05; **p < 0.01; ***p < 0.001.

Findings from archival salary data.

Along with the general observation in the factory, we examined a longitudinal dataset covering worker productivity in terms of gross salary of a 32-worker sewing group from June to December 2014. We offer several main findings from the dataset about worker productivity, which motivated the pilot intervention:

1. There are large variations in productivity across workers (refer to Figure S1 and S2).
2. There are moderate to large within-worker variations in productivity across time (see Figure S3).
3. The structural factors that the factory is responsible for, like hours, overtime pay, and work assignments, do not explain all the variations in productivity (see Table S9).
4. More experience does not entail higher productivity; in fact, experienced workers tend to produce less than newer ones (see Table S9).

There are two types of variations in productivity: within-worker and between-worker variations. Within-worker variations refer to how much one worker’s productivity varies across time (a month being a unit), whereas between-worker variations refer to how much different workers’ productivity varies during the same time period (a worker being a unit). We found sizeable variations in productivity both within and between workers. However, there is a much smaller variation in working hours of each worker. Combining with qualitative evidence, we can infer that there are considerable margins in worker productivity. In other words, despite the stable long hours each worker spends at work, they are not always achieving the level of productivity that they would like to achieve.

In Table S9, we regressed workers’ piece earnings on unit price, quantity, normal and overtime hours, and work experience, with fixed effects on the orders received by the group over the period. The structural factors do not explain most of the variations, even when a fixed-effect model tends to radically reduce total variation and exaggerate the R-square (Nickell, 1981). Thus we think for this reason there are group dynamics that can explain this variation, in particular the kind of group dynamics that Lewin studied and learned to manage.

Contrary to common beliefs of the workers and supervisors, work experience does not necessarily increase worker productivity. As the skills involved in sewing tasks are low, workers do not need extensive training before they become experienced in their jobs. As a worker reflected in an interview, she thought she did not work as hard at the time of the interview as
when she first got the job. We speculate nonstructural motivational factors may contribute to this difference.

The tremendous variations in workers’ salary earnings under similar working hours are intriguing. The variations may come from the worker herself, such as fluctuation in motivation, stress at work, or other nonstructural factors related to her perceptions of the work at hand and her group. Through the Lewinian participatory meeting intervention, we hoped to stabilize these nonstructural variations and help the workers work better and more efficient under the same effort. From the analysis of the qualitative observation and archival salary data, we see the potential of a participatory meeting intervention to increase worker productivity and change the social outcomes, including workplace empowerment like job satisfaction, sense of control, and attitudes toward participatory work and group life, which further stabilizes a high level of productivity.
Figure S1. Gross salary variation across different workers in the month of July 2014. There were 30 active workers earning wages in July. The x-axis indicates each worker and the y-axis indicates gross salary in CNY (Chinese Yuan).
Figure S2. Hourly pay variation across different workers in the month of July 2014. There were 30 active workers earning wages in July. The x-axis indicates each worker and the y-axis indicates hourly pay in CNY (calculated from gross salary divided by total working hours in July). Note: worker 28 was paid by a fixed rate rather than piece rate; thus her hour pay was counted as 0.
Figure S3. Gross salary variation plotted across time for each worker on one 32-worker group, from June to December 2014. The x-axis indicates months and the y-axis indicates gross salary. Note: the incomplete plots were either for workers whose wage did not depend on individual productivity (e.g., group leaders and fixed-wage workers), or workers who took a leave or left the job during this period.
Table S9. Panel data regression table of piece earnings on unit price, quantity, working hours, and working experience, with order fixed effects. There were 58 exogenous orders received over the time period. The structural factors that the factory is responsible for, like hours, overtime pay, and work assignments, do not explain most of the variations, indicating the role of social psychological factors in determining worker productivity.

<table>
<thead>
<tr>
<th>Piece Earnings</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Price</td>
<td>25.67**</td>
<td>(.369)</td>
</tr>
<tr>
<td>Quantity</td>
<td>1.27**</td>
<td>(.018)</td>
</tr>
<tr>
<td>Normal hours</td>
<td>1.62**</td>
<td>(.141)</td>
</tr>
<tr>
<td>Overtime hours</td>
<td>.31</td>
<td>(.227)</td>
</tr>
<tr>
<td>Work experience</td>
<td>-1.09**</td>
<td>(.176)</td>
</tr>
<tr>
<td>Order1 FE</td>
<td>15.77</td>
<td>(27.28)</td>
</tr>
<tr>
<td>Order2 FE</td>
<td>-38.43</td>
<td>(27.58)</td>
</tr>
<tr>
<td>Order58 FE</td>
<td>-75.33+</td>
<td>(43.06)</td>
</tr>
<tr>
<td>Constant</td>
<td>-48.81+</td>
<td>(27.27)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>7,606</td>
<td></td>
</tr>
</tbody>
</table>

+ p<0.1;  * p<0.05;  ** p<0.01
Table S10. Mean and standard deviations for RA reports for the six participatory meetings in the main experiment.

<table>
<thead>
<tr>
<th></th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Week 6</th>
<th>Six week ave.</th>
</tr>
</thead>
<tbody>
<tr>
<td>% time spent in problem solving</td>
<td>22.75</td>
<td>35.13</td>
<td>44.97</td>
<td>42.84</td>
<td>60.33</td>
<td>56.69</td>
<td>43.49</td>
</tr>
<tr>
<td></td>
<td>(21.06)</td>
<td>(20.74)</td>
<td>(17.90)</td>
<td>(17.52)</td>
<td>(17.86)</td>
<td>(21.59)</td>
<td>(22.97)</td>
</tr>
<tr>
<td>Problem solving rating (1-4)</td>
<td>1.83</td>
<td>2.13</td>
<td>2.23</td>
<td>2.52</td>
<td>3.73</td>
<td>3.38</td>
<td>2.62</td>
</tr>
<tr>
<td></td>
<td>(0.91)</td>
<td>(0.63)</td>
<td>(0.63)</td>
<td>(0.81)</td>
<td>(1.28)</td>
<td>(1.20)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>% time spent in general discussion</td>
<td>29.22</td>
<td>21.20</td>
<td>24.97</td>
<td>20.03</td>
<td>17.87</td>
<td>18.21</td>
<td>21.99</td>
</tr>
<tr>
<td></td>
<td>(18.68)</td>
<td>(14.88)</td>
<td>(11.96)</td>
<td>(10.82)</td>
<td>(9.77)</td>
<td>(11.02)</td>
<td>(13.66)</td>
</tr>
<tr>
<td>General work voice rating (1-4)</td>
<td>1.70</td>
<td>1.97</td>
<td>1.93</td>
<td>1.84</td>
<td>1.97</td>
<td>1.54</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.72)</td>
<td>(0.69)</td>
<td>(0.78)</td>
<td>(0.76)</td>
<td>(0.58)</td>
<td>(0.72)</td>
</tr>
<tr>
<td>% time spent in nonproduction discussion</td>
<td>11.90</td>
<td>11.20</td>
<td>4.17</td>
<td>5.77</td>
<td>3.23</td>
<td>5.92</td>
<td>7.05</td>
</tr>
<tr>
<td></td>
<td>(13.40)</td>
<td>(16.54)</td>
<td>(8.21)</td>
<td>(9.83)</td>
<td>(5.69)</td>
<td>(11.96)</td>
<td>(11.80)</td>
</tr>
<tr>
<td>Nonwork voice rating (1-4)</td>
<td>1.87</td>
<td>2.03</td>
<td>1.63</td>
<td>1.71</td>
<td>2.03</td>
<td>1.46</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>(1.41)</td>
<td>(1.13)</td>
<td>(0.93)</td>
<td>(0.97)</td>
<td>(1.22)</td>
<td>(0.65)</td>
<td>(1.09)</td>
</tr>
<tr>
<td>Interference (1-3)</td>
<td>2.00</td>
<td>1.65</td>
<td>1.63</td>
<td>1.48</td>
<td>1.48</td>
<td>1.71</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>(0.82)</td>
<td>(0.63)</td>
<td>(0.72)</td>
<td>(0.51)</td>
<td>(0.64)</td>
<td>(0.62)</td>
<td>(0.67)</td>
</tr>
<tr>
<td>Scold (1-3)</td>
<td>1.28</td>
<td>1.50</td>
<td>1.33</td>
<td>1.21</td>
<td>1.33</td>
<td>1.29</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>(0.54)</td>
<td>(0.65)</td>
<td>(0.55)</td>
<td>(0.41)</td>
<td>(0.55)</td>
<td>(0.55)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>Encouragement (1-3)</td>
<td>1.53</td>
<td>1.23</td>
<td>1.27</td>
<td>1.10</td>
<td>1.13</td>
<td>1.04</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
<td>(0.43)</td>
<td>(0.52)</td>
<td>(0.30)</td>
<td>(0.43)</td>
<td>(0.20)</td>
<td>(0.49)</td>
</tr>
</tbody>
</table>