Manager Communication Style, Worker Stress and Productivity in Bangladeshi Garment Factories* (Pre-Analysis Plan)

Paula Lopez-Pena†
(IPA & Yale University)

Kamruzzaman Mozumder
(University of Dhaka)

Atonu Rabbani
(University of Dhaka)

Christopher Woodruff
(University of Oxford)

July 24, 2018

Abstract

We conduct a randomized controlled trial (RCT) to evaluate the impact of on-the-job soft skills training for supervisors on line productivity, worker attendance, and psychological well-being in five garment factories in Bangladesh. Line supervisors and chiefs are randomly selected to receive either ten sessions of cognitive behavioral therapy featuring effective communication and stress management techniques; ten health information sessions (active control intervention), or no intervention (pure control group). This document outlines the outcome variables and econometric methods that will be used to assess the effect of training on productivity-related outcomes, worker attendance and retention, and psychological well-being among supervisors and employees.

*We thankfully acknowledge funding from the International Growth Centre and J-PAL's Urban Services Initiative. Human Subjects Approval was obtained from Innovations for Poverty Action (IPA IRB Protocol 10197B). All views expressed and errors are our own.

†Corresponding author: plopez-pena@poverty-action.org

1 Introduction

The importance of managerial style in explaining differences in productivity, profitability, and growth across firms is increasingly being recognized (Bertrand and Schoar (2003); Bloom and Van Reenen (2007); Bloom et al. (2012); Bloom et al. (2013); Bloom et al. (2016); Bruhn et al. (2018)). Various management practices, including employee motivation techniques, have long been believed to affect worker retention and productivity (Ichniowski et al. (1997); Nagin et al. (2002)). In particular, there is evidence to suggest that employee productivity increases when supervisors adopt a less authoritative style, such as performing front-line production tasks alongside workers when problems arise (Babbitt (2016); Ranganathan and Shivaram (2017)). Nonetheless, the extent to which the effectiveness of specific practices is contingent on the situation is not yet fully understood (Ranganathan and Shivaram (2016); Ayman et al. (2009)).

In this paper we ask whether offering training in effective communication to line supervisors and chiefs in Bangladeshi garment factories affects line productivity, attendance, and retention. Because previous studies suggest that high levels of work pressure and supervisor stress are an important cause of inadequate communication styles, such as using verbal abuse to motivate workers (ILO (2016)), the training combines modules on assertiveness, time, and stress management. We recruit 255 line supervisors and chiefs and randomly allocate one-third to receive ten sessions of training. Another one-third were allocated to an active control group and received ten informational sessions on general health and well-being. The remaining were kept as a pure control group. The inclusion of an active control intervention allows to disentangle the effect of teaching effective communication skills from that of receiving professional attention and information on how to increase in well-being levels.

2 Interventions

2.1 Training in Stress Management and Communication Skills (Cognitive Behavioral Therapy)

The training is designed to reduce psychological stress among line supervisors and chiefs and promote a less aggressive communication style, even when work pressure is high. It emphasizes the benefits for supervisors and employees of adopting a non-aggressive demeanor and teaches strategies to manage stress effectively, including time and anger management techniques.

The topics covered in the sessions include: (i) Education on the importance of stress management and interpersonal effectiveness, and their impact on health and job performance; (ii) Interpersonal effectiveness (includes emotional regulation and effective responses); (iii) Stress management and attention training (mindfulness and breathing techniques); (iv) Activity scheduling and time management; (v) Problem solving and alternative thinking (cognitive restructuring)

A team of clinical psychologists from Dhaka University guided the recruitment of psychologists and the selection of topics, materials and procedures to be used during the training sessions. We conducted in-depth interviews with a small sample of supervisors and line chiefs to identify important sources of stress and the relevance of the topics for each session.

2.2 Informational Health Intervention

The active control intervention consists of 10 individual, face-to-face sessions with a health worker, in which participants receive handouts and brochures with general health information. The contents of the handouts are discussed, but no specific advice or training on how to develop new skills be provided. The handouts include tips to maintain a healthy diet on a budget, and the importance of sleep and physical exercise to improve health and cognitive performance.

2.3 Pure Control

Participants in the control group take part in baseline and follow-up surveys, but do not receive any treatment.

3 Evaluation Questions

Our main research questions are:

1. Does teaching line chiefs and supervisors effective communication and stress-management strategies improve productivity and psychological well-being among supervisors and employees?

2. Is this intervention cost-effective for employers?

We also aim to provide evidence on the following:

1. What are the mechanisms through which the training has an effect on well-being, attendance, and production outcomes? Is the effect of the intervention caused by the expectation that well-being levels among supervisors will increase?

2. Is training more effective for women? Or is the effectiveness of male and female trainees contingent on the characteristics of the lines they supervise (e.g. gender composition of employees, product complexity)?

3. Is the effect of training on productivity and well-being levels more noticeable when work pressure is high?

4 Evaluation Design

4.1 Sampling Strategy

We use a Randomized Control Trial (RCT) to establish a causal relationship between the two treatment arms and changes in outcomes. We were interested in training workers in supervisory roles. We randomly recruited 255 line supervisors (LS) and chiefs (LC) from five factories in Dhaka from a roster of 355 candidates nominated by factories. In each factory, we randomly allocated one-third of the LC/LS to one of the two treatment conditions, or to a pure control group. Of the 255 participants, 184 were male and 41 were female.

4.2 Outcome Variables

Below we list the outcome variables and indices which we will consider:

1. Productivity. Factories record daily output (number of pieces produced) for each sewing line and the Standard Minute Value (SMV) of the garment being produced. The SMV is the amount of time that industrial engineers estimate it would take to produce one unit of the specified garment by a fully efficient production line; this measure allows to compare the efficiency of lines producing different garments. Productivity is estimated as:

$$Productivity = \frac{Output * SMV}{(Operators + Helpers) * Hours * 60} \tag{1}$$

2. Alteration rates. Factories record the number of garments that require some re-work and the number that is be rejected

3. Absenteeism and employee retention, based on factory records

4. Supervisor and employee stress levels: both self-reported (survey data) and biomarkers for stress (hair cortisol). Self-reported stress will be measured using an aggregate index of symptoms of anxiety (GAD-7 score) and depression (PHQ-9 score)

5. Management style: includes use of praise or abusive language as a means to motivate workers, and engagement of supervisors in front-line production tasks, as described by workers (survey data)

4.3 Data Collection Methods and Instruments

Productivity, absenteeism, and alteration rates are recorded daily by factories, and collected on a monthly basis by the research team. The data collection starts 2 to 3 months prior to the beginning of training, and continues for at least 6 months after the end of the training.

We collect self-reported data on symptoms and sources of stress, anxiety and depression, among trainees and employees under their supervision. Trainees are interviewed in person before the first session takes place (baseline), after the training is completed, and 6 months later. Follow-up surveys take place over the phone.

We measure self-reported stress using a combination of self-assessment scales - namely, the GAD-7 and PHQ-9. The GAD-7 is a 7-item form that measures the intensity of generalized anxiety disorders and related symptoms (Spitzer et al. (2006); Löwe et al. (2008)). The PHQ-9 has been validated and extensively employed as a severity measure of depressive symptoms (Kroenke et al. (2001); Moriarty et al. (2015)), with recent evidence showing its appropriateness to capture changes over time (Löwe et al. (2004)).

We measure the levels of stress hormone cortisol among supervisors, a biomarker for stress, by collecting one round of hair sample after receiving the treatment. Hair cortisol has emerged as a promising method to measure exposure to systemic stress over longer periods of time (Russell et al. (2012)). Donating a hair sample is voluntary. Hair samples are at least 1 centimeter long. The analysis is carried out by the department of Genetic Engineering and Biotechnology of the University of Dhaka.

Supervisors are also asked questions about basic individual and households characteristics, personality traits (Big Five Personality Inventory), attitudes towards learning and beliefs about their ability (growth mindset, self-efficacy). Raven's Progressive Matrices and Digit Span tests are administered at the time of the baseline to measure reasoning ability and general intelligence. We also measure the body weight and height of line supervisors and chiefs to calculate their body mass index (BMI). Finally, we pilot the feasibility of using wearable technology (wristbands) to monitor sleeping hour, physical activity, heart rate and blood pressure.

In addition, we select a sample of line operators (between 4 and 6 per line) to conduct a phone-based survey after work. We ask them questions about communication practices in the factory (i.e. such as the use of praise or verbal abuse by managers to motivate workers under their supervision), and the performance of their line supervisor under pressure.

4.4 Power calculations

We originally planned a sample size of 204 participants from 4 factories and assumed an intraclass correlation of 0.7. This sample size was chosen such that we would be able to detect an increase in production efficiency of 2% with 80% power. Data on productivity (efficiency) is recorded daily by factories. We use 90 days of pre-treatment data and 120 days of post-treatment data. We assume a correlation between baseline measurements and baseline and follow-up measurements of 0.7.

5 Econometric Specifications

5.1 Basic specifications

We begin our analysis by exploring baseline balance between each of the two treatment groups and the pure control group by regressing each baseline characteristic on the two treatment dummies.

To assess the effect of each treatment on primary and secondary outcomes of interest we use an ANCOVA specification of the form (McKenzie (2012)):

$$Y_{it} = \beta_0 + \beta_1 * T_i + \beta_2 * AC_i + \beta_3 * Y_{i,t=0} + \beta_4 * X_{h,i} + \gamma_j + \tau_t + \epsilon_{it} \tag{2}$$

where Y_{it} is an outcome for supervisor i in period t (for example, line productivity, quality defects, or absenteeism rates), T_{i} is the treatment dummy and takes value 1 if the owner was assigned to training in communication skills, AC_{i} is the active control dummy, and β_1 is the impact of training. $Y_{\mathrm{i,t=0}}$ is the lagged outcome variable, $X_{h,i}$ controls for baseline covariates, γ_j is a factory/week fixed effect, and

τ_t is a line fixed effect. Standard errors are clustered at the line level when data from different survey rounds are pooled to increase power. Robust standard errors are used in cross-sectional analyses.

In the case that more than one line is assigned to the same supervisor, all lines supervised by that individual will be allocated to the same experimental group. Because attendance and compliance with treatment protocols is voluntary, these are intent-to-treat estimates.

5.2 Heterogeneous Treatment Effects

We will test whether the impact of each treatment varies with pre-determined household and individual characteristics, measured at baseline and denoted by X_h, by estimating a regression of the form:

$$Y_{it} = \beta_0 + \beta_1 * T_i + \beta_2 * T_i * X_{h,i} + \beta_3 * X_{h,i} + \beta_4 Y_{i,t=0} + \tau_t + \epsilon_{it} \tag{3}$$

where Y_{it} is an outcome (e.g. stress index, efficiency, or attendance) for line i in period t, $Treat_{\text{i}}$ is the treatment dummy, $X_{h,i}$ controls for dimensions of heterogeneity. $Y_{\text{i,t=0}}$ is the baseline outcome variable. To address concerns over multiple hypothesis testing, we will report Sidak-adjusted p-values. As before, standard errors will be clustered at the line level. The dimensions of heterogeneous effects are:

1. Supervisor gender
2. Supervisor age and tenure. We will test whether younger or less experienced supervisors are more open to changing their managerial style, and whether workers respond differently when they do
3. Supervisor educational level. We will test whether training is more effective for trainees with a higher educational level
4. Supervisor openness to learning, measured using a combination of personality traits (big five personality inventory) and attitudes towards learning (growth mindset)

5. Baseline cognitive function, measured by Raven's matrix and digit span score
6. Baseline psychological well-being, using both self-reported measures and hair cortisol levels (high versus low)
7. Gender composition of the line supervised by trainee, using both the presence (binary outcome) and fraction of male operators
8. Product characteristics (e.g. complexity, SMV)
9. Supervisor-employee personality match. Using interaction effects, we will investigate whether employees scoring high in certain traits, such as agreeableness and conscientiousness, respond better or worse to trainees, both on average, by gender and by manager personality traits

5.3 Accounting for multiple inference

We employ three approaches to address concerns over multiple hypothesis testing:

1. We will limit our analysis to the outcomes and dimensions of heterogeneity of interest specified in this pre-analysis plan, and we will report p-values adjusted for multiple hypotheses testing
2. We will look at standardized treatment effects within domains by aggregating outcomes for which the hypothesized treatment effect goes in the same direction (Kling et al. (2007)). We will standardize each variable as a z-score by subtracting the control group mean and dividing by the control group standard deviation
3. We will follow the approach suggested by (Fink et al. (2014)) to control for multiple hypothesis testing when analyzing heterogeneity in treatment effects

5.4 Survey Attrition

Response rates may vary over time among supervisors and line operators across treatment, active control, and pure control groups. We will use probability weights to account for attrition. First, we will determine whether attrition is correlated with assignment to treatment using a regression of the form:

$$A_t = \beta_0 + \beta_1 * T_i + X_s \delta + \epsilon_i \tag{4}$$

Where A_i be a dummy variable that takes value 1 if worker i attrits from the study, and X_s are dummy variables for each randomization strata s. If treatment status and attrition are not correlated (β_1 is not significantly different from zero at the 5 percent significance level) then our estimations will not be adjustment for attrition. If attrition is found to be correlated with treatment status, we hypothesize that attrition will be higher for the active control and pure control groups. We will then use Lee's bounds to test for robustness to attrition (Lee (2009)).

6 Informed Consent and Ethics Approval

We obtained ethics approval from the Innovations for Poverty Action (IPA IRB Protocol 10197B). Informed consent to participate in the surveys, provide a sample of hair, and receive either 10 sessions of training or health information was obtained before the start of the baseline survey. The survey included a question at the beginning of the questionnaire requiring the enumerator to confirm that informed consent was obtained.

References

Ayman, R., Korabik, K., and Morris, S. (2009). Is transformational leadership always perceived as effective? Male subordinates' devaluation of female transformational leaders. *Journal of Applied Social Psychology*, 39(4):852–879.

Babbitt, L. (2016). Supervisory skills training. Impact evaluation. Technical report, International Labour Office. Geneva: ILO.

Bertrand, M. and Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, 118(4):1169–1208.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1):1–51.

Bloom, N., Genakos, C., Sadun, R., and Van Reenen, J. (2012). Management practices across firms and countries. *The Academy of Management Perspectives*, 26(1):12–33.

Bloom, N., Sadun, R., and Van Reenen, J. (2016). Management as a technology? Technical report, National Bureau of Economic Research.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, 122(4):1351–1408.

Bruhn, M., Karlan, D., and Schoar, A. (2018). The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico. *Journal of Political Economy*, 126(2):635–687.

Fink, G., McConnell, M., and Vollmer, S. (2014). Testing for heterogeneous treatment effects in experimental data: False discovery risks and correction procedures. *Journal of Development Effectiveness*, 6(1):44–57.

Ichniowski, C., Shaw, K., and Prennushi, G. (1997). The effects of human resource management practices on productivity: A study of steel finishing lines. *The American Economic Review*, pages 291–313.

ILO (2016). Progress and potental: How Better Work is improving garment workers' lives and boosting factory competitiveness: A summary of an independent assessment of the Better Work programme. Technical report, International Labour Organization. Geneva: ILO.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kroenke, K., Spitzer, R. L., and Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9):606–613.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., and Herzberg, P. Y. (2008). Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Medical care*, 46(3):266–274.

Löwe, B., Kroenke, K., Herzog, W., and Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: sensitivity to change of the patient health questionnaire (PHQ-9). *Journal of Affective Disorders*, 81(1):61–66.

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2):210–221.

Moriarty, A. S., Gilbody, S., McMillan, D., and Manea, L. (2015). Screening and case finding for major depressive disorder using the patient health questionnaire (PHQ-9): A meta-analysis. *General hospital psychiatry*, 37(6):567–576.

Nagin, D. S., Rebitzer, J. B., Sanders, S., and Taylor, L. J. (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review*, 92(4):850–873.

Ranganathan, A. and Shivaram, R. (2016). (Managerial) style over substance: Determinants of devaluation for female supervisors in an Indian garment factory. *Working paper*.

Ranganathan, A. and Shivaram, R. (2017). Leading by doing: How female supervisors motivate worker productivity through subordinate scut work. *Working paper*.

Russell, E., Koren, G., Rieder, M., and Van Uum, S. (2012). Hair cortisol as a biological marker of chronic stress: Current status, future directions and unanswered questions. *Psychoneuroendocrinology*, 37(5):589–601.

Spitzer, R. L., Kroenke, K., Williams, J. B., and Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD7. *Archives of Internal Medicine*, 166(10):1092–1097.