# Evaluation of the Work-Progression & Productivity Toolkit

**FINAL REPORT**

**January 2018**

**Christopher Woodruff**

**Hannah Uckat**

**Anaise Williams**

This report is based on data from the third phase of a project that began in the fall of 2011. We are grateful for the cooperation of many partners who made this research possible. We thank first Better Work Bangladesh for providing many factories that provided access to their employees for surveys, and access to very detailed and confidential production data. We are grateful to World Bank and IFC that provided both funding and other support. Innovations for Poverty Action (IPA) coordinated the survey and data gathering work, and the efforts of Leigha Miyata, Anaise Williams, Suraia Akther, Fazle Ellahi, Abu Taher Mithu, Mir Rafiat Salman, Jesmin Sultana, Abdul Awal, Shankar Dey, Marwan Rashid, Mahmudur Rahman, Nutan Farah Haq and Quazi Farhan are greatly appreciated. We also offer thanks to IPA trainers - Sohel Kabir and Rumana Afroj Mukti for designing and delivering technical training, selection training, on-boarding training and team building session to factory workers and management. Also we offer thanks to Ashna Chowdhury of Thrive and Shipra Chowdhury and Maruf Hassan Khan of Better Work Bangladesh for carrying out combined soft skill training to workers. We extend special thanks to Eleonore Richardson and Nabeera Rahman of IFC; Louis Vanegas, Nayeem Imran, Kesava Murali, Shamsul Haque, Saroar Hossen and all Enterprise Advisors of Better Work Bangladesh for support with carrying out the project. Responsibility for any errors lies entirely with the Principal Investigators.

Principal Investigators:

Christopher Woodruff, Professor of Economics, Department of International Development, University of Oxford, Oxford, UK.

Hannah Uckat, Ph.D. Candidate, Department of International Development, University of Oxford, Oxford, UK.

# Executive Summary

This report summarizes findings from a randomized control trial evaluating the Work-progression & Productivity Toolkit (WPT). The goal of WPT is to provide female operators with skills that lead to promotion to supervisory positions. The Bangladeshi garment industry was instrumental in expanding women's employment opportunities in the urban job market. However, the positions available to women remain mostly limited to those at the machine-operator level. Social norms around supervisors being male continue to limit female career aspirations. Efforts to encourage promotion of female workers can advance both gender empowerment and industry development.

To measure the impact of WPT, RMGPP draws from survey data with trainees, operators, line chiefs and line supervisors, and upper management and from factory administrative data from a sample of 27 participating factories. To isolate the effectiveness of training, we randomized individual workers into three groups (soft skills training, soft and hard skills training, and control). We seek to determine whether soft skills training or hard skills training is a more effective and efficient means to promoting female advancement.

The following are some highlights of program impact:

1) Out of the 134 program completers who were surveyed in January 2017, 74 (or 55%) are currently working as a supervisor or assistant supervisor. Of the 10 factories who had 0 female supervisors at the start of the program, 9 factories offered promotions to trainees and 8 had at least one trainee accept a promotion.
2) Baseline diagnostics help identify candidates more likely to be promoted to supervisor. Breaking diagnostics into those related to aptitude and those related to attitude, we find that trainee attitude is particularly important. Baseline attitude scores are positively associated with higher training completion rate, higher promotion rates, more confidence in the role of supervisor, and higher performance ratings by operators being supervised, and a more cooperative supervisory style
3) We find evidence that the training programs led to higher levels of participation in the program and that the hard and soft skills training increases trainee confidence across ten main skill areas that supervisors should have to be successful at their job.

4) Operators report that trainees have a more cooperative supervisory style than typical supervisors. We found that trainees rated more highly by operators are more likely to be offered promotion and more likely to accept a promotion.

While these findings are promising, we also found some limitations with the program. First, the program dropout rate was high. Second, adherence to the initial agreement that trainees would work daily as assistant supervisors for eight weeks was lower than expected. Even the trainees that completed the program worked only about two-thirds of the agreed number of days as assistant supervisors. Third, while over half of participants were offered a promotion within the months following the end of the training, there was a surprisingly low rate of acceptance. Looking at self-reported reasons for dropping out or not accepting an offer of promotion, the majority of responses centered around not liking the content and responsibilities of the supervisor role. Fourth, while the combined soft and hard skills training appears to have increased the confidence of trainees in the ability to perform in the supervisory role, there was little change in overall measures of garments knowledge from before to after the training across treatment groups. Finally, we are not yet able to show impacts of the training on the efficiency of the trainees. This reflects both the complexity of the administrative production data and far from perfect compliance with protocols around assigning trainees to specific trial lines.

A common theme in our results is that aptitude is most strongly associated with being offered a promotion, while attitude, particularly family support, is more important in ensuring both that a trainee stays in the program and accepts an offer for promotion. Given this in combination with our other findings, we present the following recommendations:

1) When selecting training participants, management should select nominees who express a strong interest in the position, and have higher levels of confidence and strong family support for becoming a supervisor. Higher years of schooling is also predictive of completing the program.
2) To ensure higher program impact on skills, it is important to ensure that trainees get to on-the-job practice through trialing as assistant supervisors. In this program, trainees only trailed about 60% of the time they were supposed to, which was a large limitation on their ability to make impact on productivity.
3) Soft skills training alone appears not to be effective. The training in soft and hard skills combined is effective in increasing the confidence level of trainees. We cannot say for certain whether the larger effect of the combined training results from the diversity of skills being taught or simply the increased training time. But the evidence supports a longer training period, and the combination of soft and hard skills has important impacts.
4) Finally, our results demonstrate that norms characterizing females as less able to be supervisors remain prevalent among different levels of factory workers. These norms likely discourage female

willingness to be promoted. The selection training offered to middle- and upper-level managers included a discussion of norms and implicit biases. The training was effective in getting managers to select trainees with characteristics associated with higher levels of promotion and effectiveness as a supervisor.

# Table of Contents

# Table of Figures

# Project Overview

The Bangladeshi garment industry was instrumental in expanding women's employment opportunities in the urban job market. After nearly 40 years of industry development, women make up the majority of production floor workers, but only a tiny share of managers in the industry. Social norms around supervisors being male continue to limit female career aspirations, with women's advancement remaining capped at the level of a highly skilled machine-operator. As the opportunities for men continue to expand outside the garment sector, efforts to encourage promotion of female workers are important for gender empowerment and industry development.

The Work-progression & Productivity Toolkit (WPT) is a series of training interventions designed to provide female operators with skills that prepare them for promotion to supervisory positions. In addition to improved gender equality in the workplace, WPT has the potential to improve productivity and quality in factories. The majority of sewing section workers is women and factories are more likely to find suitable candidates for promotion by widening the pool of candidates to include women.


RMGPP is working to evaluate whether WPT is effective. Training is conducted at two levels. First, in half of the factories, senior managers are trained in new methods for selecting women with the potential to be supervisors. A series of diagnostics measuring both aptitude and attitude are provided to managers. Second, training is provided to female sewing machine operators selected by all of the factories. Evaluation of the effect of this training is aided by randomizing individual workers into three groups (soft skills training, soft and hard skills training, and control). We seek to determine whether soft skills training or hard skills training is a more effective and efficient means to promoting female advancement. Participating factories committed to trialing all trainees, including the control group, as assistant supervisors for at least 8 weeks. To measure the impact of WPT, RMGPP worked with local partner Innovations for Poverty Action (IPA) to collect both survey data and factory administrative data. This report outlines the final findings from the both sources of data.

## Project Stages

Together with BetterWork, we collected a cohort of 27 factories to participate in the project. Innovations for Poverty Action asked each factory to nominate a number of female operators to participate in the Work-progression & Productivity Toolkit program. The number of

operators to be trained varied by factory according to the expected needs for new supervisors. The factories were asked to rank all potential trainees, and to provide a number of replacements in case those in the highest priority group were unable or unwilling to take part in the training. The factories returned a nomination form, after which a first meeting with the research and implementation team was scheduled. At the first meeting, IPA explained the entire project, including commitments, and the cooperation needed from the factory. IPA visited the factory again 1-2 weeks later to conduct the baseline survey and first collection of production data. Data collection continued throughout the program, with IPA staff visiting each factory every three months.

Half of the factories were randomly selected to receive training on selecting appropriate candidates for training. This training was aimed at mid- to high-level managers, those involved in nominating the female operators attended the Selection Training. Technical trainers from IPA gave training on using quantitative skills and attitude assessments along with “soft” knowledge of potential trainees. After the training, factories were invited to revise the initial list of trainees provided to the project team. On the same day, a survey team from IPA gave a variety of tests to the nominees to screen qualified trainees. The trainees were then allocated by lottery to one of three equally-sized groups that received training at different times.

At the beginning of the training program, we conducted and “on-boarding” session. The idea was to introduce the new trainees to the lines on which they would work. Participation by high-level management was intended to signal the importance of the project. Managers in the production department, including line chiefs and supervisors from the lines where the trainees were to trial as assistant supervisors (trial lines), and those in the worker participation committee, were invited to the on-boarding training. A high-level manager, ideally the managing director from the factory, introduced the selected trainees to everyone. One week before the training of the first group, all trainees from all three groups began working as an assistant supervisor on a pre-selected (trial) line, shadowing an existing supervisor.

Trainees in the soft and soft and hard treatment groups then began receiving classroom training. They attended classes at the training center two days a week, and continued to work as assistant supervisors on the same trial line for the remaining four work days. This continued for two to five weeks until the training was complete.

At the end of the classroom training, all trainees, supervisors, line chiefs, and IE officers in charge of the respective trial lines participated in a team-building session at the training center. This was designed to

increase collaboration among those working on the same line to ease the incorporation of the trainee in the management team.

Approximately four months after training began, and two months after the official eight-week trial period ended, a survey team visited the factory again to survey the trainees, line managers, and selected workers on the trial lines. Trainees allocated to the control (comparison) group, and those allocated to receive delayed training in hard skills, received training at the training centers immediately after this follow-up survey was completed. All trainees were thus allocated to received nine days of training outside the factory over the entire length of the program.

We continued to collect productivity data from the factory for at least six months after the last training session was complete. During this one-year project, the IPA data collection team visited each factory six times, and the IPA survey team visited each factory three times. We also conducted further short follow-up surveys with trainees by telephone to collect information on promotion.

## Intervention

Factories were divided into 5 sessions of 5 to 6 factories each for training. The first training session began in November 2016 and the fifth began in March 2017. The initial training for the Hard and Soft and Soft Only groups finished in May 2017. Training began for the control groups in August 2017, and all training was completed in October 2017. For the full training schedule, please reference Appendix 6. The following table outlines the number of trainees who fully participated in the training by trial arm:

**Table 1: Trainee Completion Status**

| Session | Hard & Soft | Soft Only | Control | Total |
|---|---|---|---|---|
| 1 | 12 | 12 | 8 | 32 |
| 2 | 9 | 9 | 8 | 26 |
| 3 | 8 | 11 | 4 | 23 |
| 4 | 10 | 11 | 4 | 25 |
| 5 | 16 | 17 | 5 | 38 |
| Total | 55 | 61 | 29 | 144 |

Note that while the same number of nominees was allocated to each treatment group, the number of trainees in the control group who received training was much lower than in either of the other treatment groups. Twenty-two trainees assigned to the control participated in the 8-week trial, but decided they did not want to be a line supervisor before their classroom training sessions started. Given the delayed starting date of the control group training, a large portion of the nominees allocated to the control group had dropped out of the trial. We discuss this, and differences between dropout rates across training groups, later in the report.

The hard skills training focused on the technical side of being a supervisor. During five full days of training, the main components were production processes, sewing machines, quality control, cutting, finishing, printing, embroidery, and the general supervisory role. The overall target of the training was to lay the basic technical foundation for the female supervisor.

A second training focused on soft skills. Over four days of soft skills training, participants discussed breaking barriers, increasing self-confidence, developing leadership skills, practicing effective communication and listening, fostering a healthy working environment, and being a team player and role model for workers. The training additionally focused on understanding harassment, developing integrity and fairness, workers' rights and responsibilities, and human resources management, including types of management styles.

In addition to training our project beneficiaries, we also provided training on nominee selection to factory management in half of the randomly selected factories. The purpose of these trainings for select factories was to 1) train factory management in standardized worker evaluations and 2) ensure factory management understood and was onboard with the training and intentions of the study. Two types of trainings were provided to factory management: selection training and onboarding training. About half of the sample factories were randomly selected to receive selection training (15 in total). A total of 160 managers involved in nominating the female operators attended this training in their factories. It is expected that these managers will have better knowledge in identifying and promoting their best workers. The managers were given the opportunity to change their preferences for which women to train after receiving the training. We will test whether this resulted in selection of trainees with characteristics of trainees associated with more success as supervisors.

The onboarding training involved managers in the production department, including line chiefs and supervisors from the lines where the trainees were assigned to trial as assistant supervisors. A total of 341 lower level managers were trained on the effectiveness of female line supervisors and the best way to support female line supervisors in succeeding in their new role.

# Data Overview

This report draws from both factory administrative data and survey data. Survey data was designed to assess demographics, attitudes, and experiences from baseline to follow-up and to contextualize productivity changes and differences between trial lines as found in the factory administrative data. For information on our survey and productivity data collection protocols, please reference Appendix 3 (factory administrative data) and Appendix 5 (survey data). If a term or concept is confusing, please reference our glossary in Appendix 7. Below, we describe the factory and trainee samples that we collected data from.

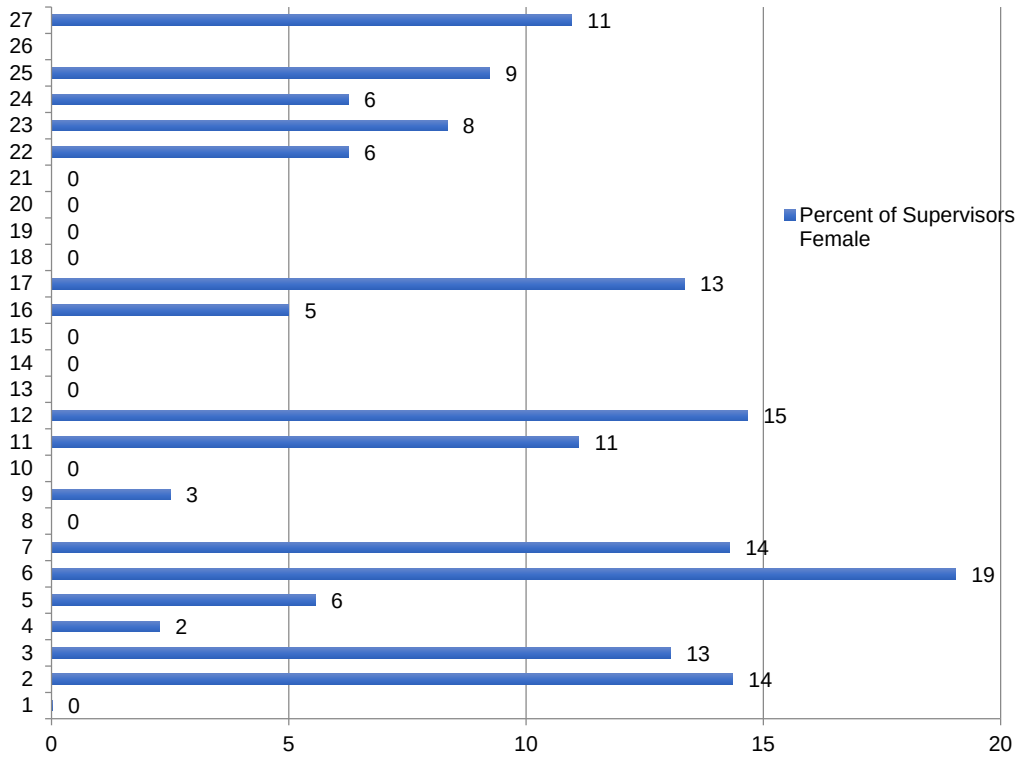
## Factory Sample

We recruited 30 factories for our study before launching in November 2016 to fulfill the number of target trainees (270, an average of 9 for each factory). BetterWork provided an initial list of 45 factories, of which 11 factories were discarded due to location and type of product produced. Four factories declined to participate. Three factories dropped out of the program before trainees were fully trained, leaving us with a sample of 27 factories.

The factories in our sample surround the city of Dhaka; 12 are in Gazipur, 3 in Narayanganj, 1 in Mirzapur, 6 in Ashulia, 1 in Mirpur, and 4 in Savar. All export to large western buyers. The firms in our sample are overall relatively large, with an average of 3,148 workers (ranging from 660 to 13,553 workers). On average, the factories have slightly more female workers than male workers. The average number of lines is 33, with the smallest factory having only 5 lines and the largest having 210.

The factories in our sample had very few female supervisors at the start of the program. The average across the factory sample is about 5%. Figure 1 shows the percentage of female supervisors at baseline in each of the factories. Ten of the 27 factories (38 percent) had no female supervisors, while in one large factory, 19% of supervisors were female.

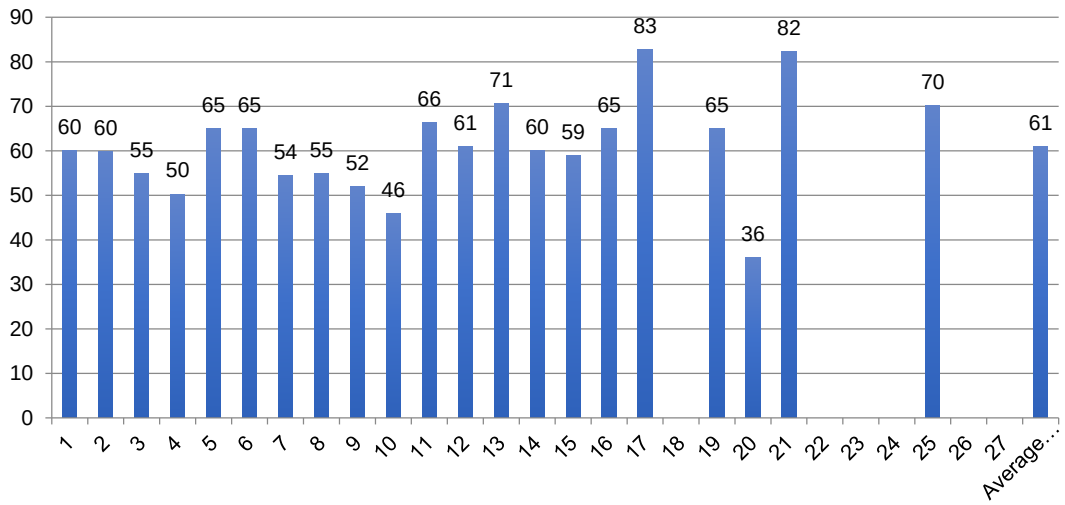
**Figure 1: Percent of Supervisors Female[1]**




On average, in our factories about 60% of sewing operators and helpers were female (figure 2).

[1] For factory 26 on this figure, the data is missing

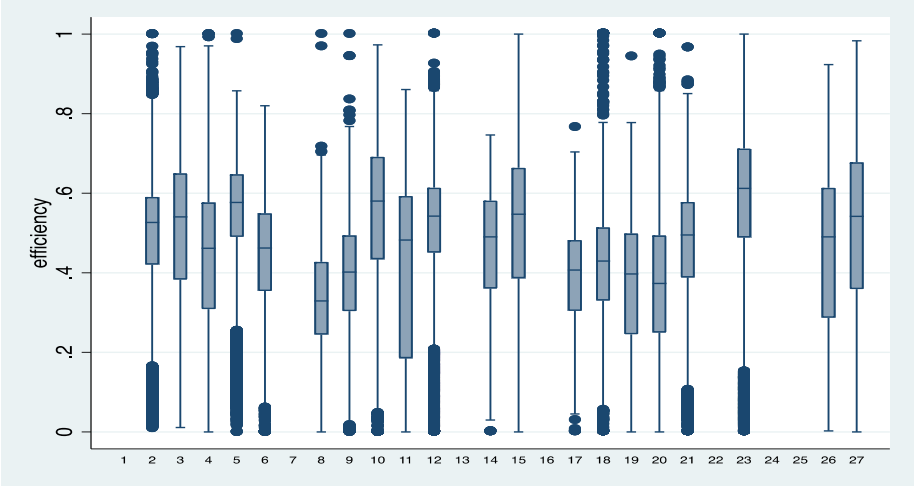
**Figure 2: Percent Operator and Helpers Who are Female by Factory at Program Start**




[2]

When looking at factory efficiency, factories show similar efficiency levels[3] of between 40%-60%. The figure below demonstrates the range in efficiency measures per WPT factory (factory identity undisclosed; each column represents a factory). Please reference Appendix 4 for more information on how we measure efficiency.

**Figure 3: Efficiency Across Factories**




[2] We are missing data on worker composition for factories 18, 22, 23, 24, 26 and 27

[3] Note that factories may adjust their standard minute value (SMV) measures to the factory context. As a result, the efficiency data may not be directly comparable across factories. We have found that the raw data will reflect the relative efficiency of a factory for most, but not all factories. Adjusting the data to a comparable base takes considerable time and effort.

At the beginning of the program, we asked upper-level management at participating factories about the process of selecting line supervisors in their factory. Over 80% of the factories said that they prefer to hire internally for supervisor positions. In the case of internal promotion, we asked managers who is consulted to gather an initial list of candidates. They gave an array of answers, but the most common response was that candidates are selected by production managers, industrial engineers, and line chiefs. The APMs, floor-in-charge, HRM, IE, and Quality Mangers are also sometimes consulted in this initial stage. The majority of the factories said that they hold a meeting to discuss the candidates on the list. HRMs have a bigger involvement here than with the list generation, while production managers, line chiefs and industrial engineers also usually attend these meetings. Generally, the factories reported that the production manager makes the final decision on which candidate to choose and confirms the promotion with HR.

When asked how long it takes for a new supervisor to be proficient, half of the factories reported that it is usually around 3 months, while 30% of the firms said it takes closer to 6 months and others said up to one year. In terms of where a new supervisor should work, 8 factories put new supervisors on the line where they work at the time, 4 on lines with only basic styles, and 12 factories said they do not have a protocol and put a new supervisor on any line. The majority of factories (19 out of 29) said that they do assign mentors to guide the newly promoted supervisors. These mentors are usually line chiefs, line supervisors, or production managers.

## Trainees

**Table 2: Trainee Demographics**

| | |
|---|---|
| Average Age | 25.7 |
| Average Years of Education | 8.4 |
| % with upper secondary education | 28.25% |
| Average Years in Garment sector | 6.3 |
| Average Years as an Operator | 5.4 |
| % Married | 75.7 |
| Average # of Children | 1.1 |

At the start of the program, the average trainee had 6.3 years of garments' experience and only 28.25% has any upper secondary education (over 9 years of education), as shown in Table 2. A key part of our baseline assessment of trainees was to conduct a series of diagnostics measuring aptitude and various aspects of attitude. For descriptions of each of the eight diagnostic measures and how we evaluated them, please reference Appendix 1. The aptitude tests included a literacy test, a numeracy test, a test of fluid intelligence (which we refer to as processing speed) and a test of garments knowledge. We also asked a series of questions intended to measure how supportive the nominee's spouse / family was of them taking on the role of supervisor, how interested in the position they were themselves, and how confident they were in their ability to perform in the role. Eligibility for the program depended on the scores in the literacy and

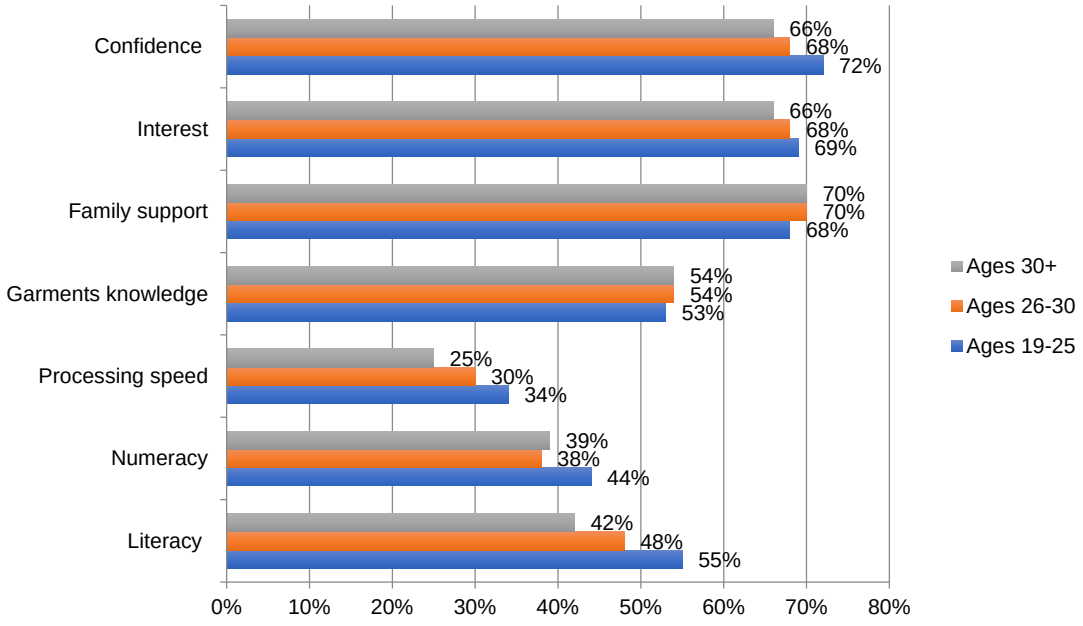
numeracy tests, but not on any of the other diagnostics. Trainees had to score above 0% on both the literacy and numeracy test, and above 25% on one of them to be eligible for the training program. Those failing these parts of the diagnostic were replaced by nominees ranked lower by the factories. Table 3 displays the scores for the various key measures at baseline across treatment groups. Generally, trainees had low numeracy and processing speed scores. Scores were similar across training groups.

**Table 3: Average Trainee Baseline Diagnostic Scores**

| | Literacy | Numeracy | Processing Speed | Garments Knowledge | Family Support | Interest | Confidence |
|---|---|---|---|---|---|---|---|
| **Overall** | **45%** | **37%** | **30%** | **52%** | **68%** | **66%** | **68%** |
| Hard + soft | 58% | 46% | 33% | 55% | 72% | 71% | 74% |
| Soft only | 56% | 48% | 31% | 54% | 74% | 79% | 73% |
| Control | 56% | 50% | 36% | 54% | 75% | 72% | 74% |

Figure 4 shows that literacy, numeracy and processing speed are highest among the youngest trainees (below age 26). Confidence was also slightly higher among the younger trainees, however, interest and family support scores were relatively the same across ages.

**Figure 4: Scoring by age**




While the differences in diagnostic scores by age are modest, there are more marked differences by years of schooling. Among the 367 nominees for whom we have data, the majority (63 percent) have exactly eight years of schooling. An additional 10 percent have seven years of schooling. Just over one-quarter, then,

have nine or more years of schooling. We find significant differences between the nominees with eight years of schooling or less, and those with nine or more years of schooling. It is perhaps not surprising that schooling is strongly correlated with scores on the aptitude portion of the diagnostic. We show this relationship on Figure 5, which groups nominees with seven or eight years of schooling, those with nine years and those with ten or more years. The graph shows the standardized average score of the literacy, numeracy, processing speed and garments knowledge parts of the diagnostic. There is a marked increase in all three aptitude scores across each of the categories, and a marked increase in garment's knowledge among nominees with 10 or more years of schooling (48, or 13 percent of the nominees).

**Figure 5: Years Schooling and Aptitude Scores**

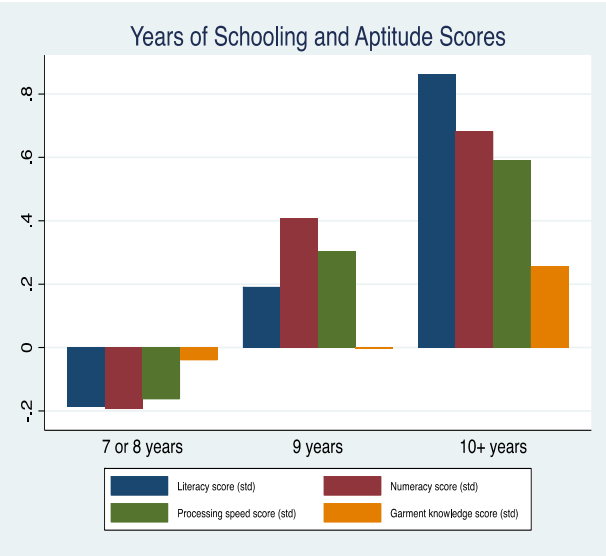
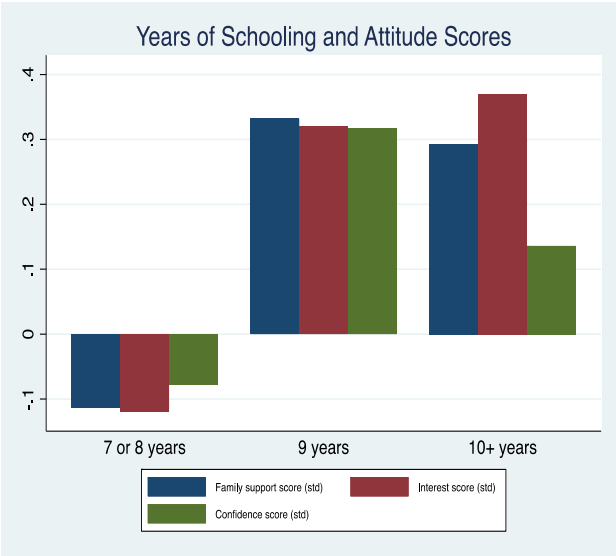



**Figure 6: Years Schooling and Attitude Scores**



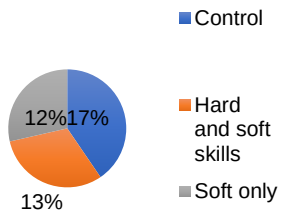

What is perhaps less obvious is that there is also a marked difference in the attitude of the nominee toward the role across education categories. These data are displayed in Figure 6. The scores on interest in being a supervisor, support from family, and confidence in performance in the role, also increase with education. But here we see the largest gap between those with eight years or less of schooling, and those with nine years or more. There is no further increase in interest and family support, and perhaps even a small drop in confidence, in moving from the group of nominees with nine years and those with ten or more years of schooling. We highlight these differences because while only a minority of workers in the factories have nine or more years of schooling, the diagnostics suggest that this more highly-educated minority may be better prepared to take on supervisory roles.

# Key Outcomes

In the following sections we cover key findings that may be useful for scale-up of WPT or similar programs. We start by discussing trends around promotions and the relatively high program dropout rate. Using the scores on diagnostics conducted with nominees, we look at whether aptitude and attitude measures predict program completion and promotion. We also look at how factories evaluate nominees, and whether a selection training with factory management has a positive impact. From there, we present findings on changes around trainee technical knowledge and confidence from before and after the trainings. We then use productivity data (which we are still in the process of collecting) to outline a preliminary evaluation of how efficiency, defect rates, and absenteeism changed on the lines that trainees worked on. Finally, we present data on operator and management personnel perceptions of trainee capabilities as well as support and resistance towards trainees working on the line as trial assistant supervisors. Based on these results, we provide recommendations for scale-up.

## Program Completion and Promotions

**Figure 7: Dropouts by treatment group**




At the start of the project, there was a large drop (34) in the number of trainees due to three factories dropping out, leaving 201 trainees across trial arm. Over the course of the eight-week trial, 56 trainees left the program. This left 145 who completed the training. This is a larger number of dropouts than expected. When looking across trial arms, there was a larger number of dropouts from the control group (figure 7).

In the weekly and follow-up surveying we attempted to contact the individuals who dropped out. Among those we reached, the main reasons for dropping out were largely due to simply not wanting the position. Many dropouts reported leaving the program because they were disinterested, while others reported family issues. A further significant reason for dropping out was leaving the factory. For trainees who left the factory for another factory, we documented this through surveying and counted offers of promotions within the new factory if they occurred.

At the time of follow-up in October 2017, there were a total of 40 offers of promotion across treatment arms (20% of the initial trainees). Of those who had completed the full training, both hard and soft skills, 22 had been promoted (33% of the hard and soft group). In January 2018, approximately 2 months after the control group training was complete, another round of surveys with trainees was conducted. In this survey, 162 trainees were reached. At this point 145 women had completed the training across trial arms, but we tried to also survey people who had dropped out of the program. We reached 162 people by phone, of which 134 had completed the program (meaning we missed 11 of the ‘completers’) and 28 had enrolled in the program but did not complete it.

Of the 162 individuals in this additional survey, 92 has been offered a promotion and 51 had accepted. Of the 92 who were offered a promotion, 86 (93%) were offered a promotion to line supervisor. One was promoted to a higher grade of operator, 4 were promoted to assistant supervisor, and 1 was promoted to sample technician. Below is a summary of promotions as of January 2018:

**Table 4: Promotions Status**

| | |
|---|---|
| Training completers offered a promotion* | 76 (56.72%) |
| Offered a promotion, total** | 92 (56.79%) |
| Number of offers to Line Supervisor or Assistant Line SV * | 74 (55.22%) |
| Training completers accepted a promotion* | 49 (64.47%) |
| Accepted a promotion, total** | 51 (55.43%) |
| Percent completers working in a higher position | 36.57% |
| Percent working in a higher position, total | 31.48% |

*n=134 (the number of trainees in this survey sample that completed the program)

**n=162 (the total number in this survey sample, which includes 28 who dropped out of the program)

The percentage of offers of promotion that are declined is both notable and of concern. Of the 145 trainees completing the training, 74 had been offered a promotion to Supervisor or Assistant Supervisor by December 2017. This ratio reflects in part the fact that factories often do not have openings at the end of the training. We have found, both the previous projects and in this one, that promotion rates continue to increase for a lengthy period after the completion of training. In that regard, perhaps a more troubling outcome is that only 49 of the 74 offers were accepted. When asked why they declined the promotion, the most common responses were related to health – combining sickness and pregnancy – given by 12 of the 25 trainees completing the program and declining a promotion. Seven gave responses indicating a lack of interest in the increased responsibility and five said their family did not support their taking the job.

In the survey, we asked people if they would want to be promoted in the future. About 70% of respondents shared that they would accept an offer in the future. Figure 8 outlines the reasons for why some people would not accept an offer. About 75% of the reported reasons for not wanting a promotion is related to not liking the content or responsibility of the job, while 5% of trainees believe managers do not want women to be promoted, and 5% of trainees themselves do not think that women should be promoted.

**Figure 9: Reasons for not accepting a promotion**

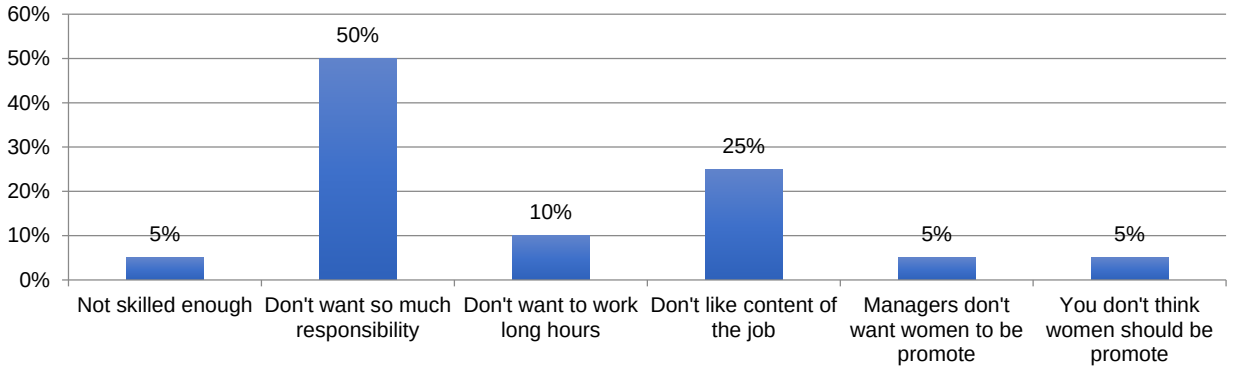



**Figure 8: How likely to be promoted in next 3 months**

When asked how long they expect to continue working in their current position, 37.5% said one year or less, 27% said between 1 and 2 years, 21.5% said between 2 and 4 years, and 14% said between 4 and 10 years. Just over a quarter of trainees said that they would certainly or very likely be promoted within the next 3 months (figure 9).

Of the group of 162 who were contacted in January 2018 survey, 19 (12%) no longer work in the garment industry at the time this data was collected. Of those no longer working in the garments industry, 18 are not working at all, of which only 6 said they have plans to return to working outside of the home. When asked why they quit, the two most common responses were that a family member fell ill and that management misbehaved. Other responses were that either the trainee or her family did not like the job, pregnancy, and childcare needs.

## Do Baseline Diagnostic Scores Predict Dropout and Promotion Outcomes?

We next ask whether the diagnostic scores predict a series of outcomes for the trainees. We consider four outcomes: 1) completion of the training program; 2) the percentage of days the trainee worked as an Assistant Supervisor; 3) whether the trainee was offered a promotion, conditional on completing the program; and 4) whether the trainee was promoted and working as a Line Supervisor in December 2017.

The last outcome reflects both an offer from the factory and acceptance of the offer by the trainee. Regressions that examine these outcomes are shown on Table 5 below.

The first two columns of Table 5 examine whether the trainee completed the training program, or dropped out before the end. The sample excludes trainees in the three factories that dropped out altogether, as we consider those to be factory-level decisions rather than trainee decisions. The regressions control for factory fixed effects, appropriate as the factory was the unit of randomization for the training treatments. Column 1 reports a probit regression, reflecting the discrete nature of the outcome. Because all of the trainees in several of the factories completed the training, these observations are dropped from the probit specification. So, we also report in the second column an OLS regression that retains these observations. We see no substantial differences between the probit and OLS regressions. The regressions suggest that attitude matters more than aptitude in determining whether a trainee completes the program. A one-standard deviation increase in the attitude score is associated with an increase of 12 percentage points in the completion rate. Aptitude and assignment to either the Soft only or Soft and Hard skills training groups all have positive measure effects, but none of these effects are statistically significant at customary levels.[4] Please reference Appendix 2 to view the same regressions with each of the individual components of the diagnostic. Notably, among the individual attitude scores, the family support score is most strongly associated with completion of the training program.

**Table 5: Outcomes: Completed Training and Promotion**

| VARIABLES | (1) Probit Completed Training | (2) OLS Completed Training | (3) OLS % Days ASV | (4) Probit Offered Promotion | (5) Probit Promoted | (6) OLS Promoted |
|---|---|---|---|---|---|---|
| Aptitude Diagnostic Score (standardized, 5 components) | 0.061 | 0.071 | -0.014 | 0.287* | 0.187 | 0.108 |
| | (0.070) | (0.063) | (0.066) | (0.157) | (0.136) | (0.100) |

[4] In an OLS regression without factory fixed effects, both aptitude and the log of factory tenure are significantly and positively associated with competing the training.

| | | | | | | |
|---|---|---|---|---|---|---|
| Attitude Diagnostic Score (standardized, 3 components) | 0.122*** | 0.123** | 0.132*** | 0.036 | 0.222*** | 0.114* |
| | (0.045) | (0.045) | (0.037) | (0.091) | (0.084) | (0.060) |
| Assigned to Soft Skill Training | 0.071 | 0.070 | 0.227*** | 0.051 | 0.151 | 0.097 |
| | (0.072) | (0.084) | (0.074) | (0.143) | (0.144) | (0.115) |
| Assigned to Soft and Hard Skills Training | 0.093 | 0.086 | 0.209*** | -0.098 | -0.006 | 0.001 |
| | (0.060) | (0.068) | (0.073) | (0.115) | (0.124) | (0.091) |
| Log age of Nominee/ Trainee | 0.174 | 0.147 | -0.314 | 0.124 | 0.247 | 0.169 |
| | (0.245) | (0.219) | (0.221) | (0.480) | (0.434) | (0.353) |
| Log factory tenure of Nominee/ Trainee | -0.007 | -0.007 | 0.011 | 0.068 | 0.064 | 0.041 |
| | (0.038) | (0.042) | (0.044) | (0.075) | (0.067) | (0.039) |
| Observations | 174 | 203 | 184 | 139 | 134 | 161 |
| R-squared | | 0.239 | 0.420 | | | 0.347 |
| Factory FE | YES | YES | YES | YES | YES | YES |

Robust standard errors in parentheses, clustered at factory level
*** p<0.01, ** p<0.05, * p<0.1

A second measure related to training is the percentage of days during the trial period that the trainee actually worked as an Assistant Supervisor. Factories agreed *ex ante* that trainees would work as ASVs all of the days they were not in training classes during an eight-week period. However, because they trainees were typically highly-skilled operators, production pressures often meant that they worked as operators during this period. Trainees that competed the training program worked as an ASV an average of 65 percent of the assigned days. We see (Column 3) that this percentage is much higher among those assigned to either of the two groups receiving immediate training. The percentage is also higher among those with higher attitude diagnostic scores. As before, we show in the Appendix 2 the regression with the individual components of the diagnostics. Here we find that the score for interest in the position has the strongest association with the outcome.

Finally, we ask if the diagnostic scores are related to whether the trainee is offered a promotion, and whether she accepts the promotion. These regressions, reported in Table 5 Columns 4 through 6, also include variables indicating assignment to one of the two groups receiving immediate training. Note, however, that the data we use comes from a follow-up phone survey conducted in December 2017, by which time all of

the trainees had received all of the training sessions. The results suggest that aptitude is most strongly associated with being offered a promotion, but attitude is more important in whether the trainee is actually working as a Line Supervisor – the combination of receiving and accepting a promotion offer. A one standard deviation increase in the attitude score is associated with an 11-percentage point increase in the likelihood of working as a Supervisor (column 5).

As described above, 12 individuals who declined a promotion gave responses relating to either not liking the job content or family resistance. This is a very small number of observations, but we nevertheless investigate whether the initial diagnostic scores flag issues of family support among the five stating family reasons and lower interest among those saying they declined out of lack of interest. We find no initial signal among those claiming family support as a reason for turning down the promotion, but the eight trainees declining for lack of interest in the position, the initial interest score was indeed significantly lower than average (-0.81 vs. 0.37 among all others completing the training, $p<.01$). This is at least somewhat reassuring with respect to validity of the initial diagnostic. The analysis suggests that the initial diagnostics are predictive of completion of the training program, being offered a promotion, and being promoted to the position of supervisor.

## Factory Selection Training

As discussed in the "intervention" section, we conducted training sessions with mid- and high-level management in 15 of the 27 factories to provide new tools for selecting potential trainees. The theory of change for the Selection Training is that the training would lead managers to select trainees better suited for supervisory positions, resulting in lower dropout rates and higher promotion rates. We will show below that the selection training did indeed result in selection of trainees achieving higher scores on our diagnostic tests. We find that re-ranking of trainees after the Selection Training increases both the aptitude and attitude scores of the trainees, with somewhat stronger effects on attitude scores.

There are two ways to test the further links on this chain. The first is to examine the Selection Training directly. That is, did trainees in factories receiving the Selection Training have higher training program completion rates and higher promotion rates? We can also examine changes in the rankings of specific trainees. Among the trainees who were above the training cutoff line both before and after the re-ranking, we can ask whether those whose ranking were more likely to complete training / be promoted than those whose ranking was decreased.

Unfortunately, with a sample of only 27 factories, we have limited statistical power for these comparisons, and we might expect the results to be only suggestive. But if we find that, for the sample as a whole,

trainees with higher aptitude and attitude scores are more likely to complete the training program and to be promoted, then the fact that Selection Training resulted in trainees with higher diagnostics scores provides some indirect evidence of a positive effect of the Selection Training.

**The Effect of Selection Training on Selection of Trainees**

We first ask whether the Selection Training induced factories to alter their initial rakings of the nominees. Because each factory selected the number of nominees based on expected need for new supervisors, the number of trainees, and hence the number of nominees, varied widely across factories. The number rank of a nominee is then not reflective of her expected suitability for the role. A nominee ranked third out of 20 may be expected to be a much better match for the position than a nominee ranked third out of four. We therefor translate all of the nomination ranks into percentiles. So, in the example just given, the first nominee would be the 85$^{th}$ percentile, while the second would be the 25$^{th}$ percentile. For each nominee that passed the literacy / numeracy diagnostic, we then take the difference between the percentile rank after Selection Training and the initial percentile rank as an indication of the degree to which the selection training affected the ranking (positively or negatively) of that nominee.

The changes in rankings following the Selection Training were substantial. Out of 150 nominees in the Selection Training factories, 25 moved from level C, the second substitute group, before the re-ranking, to Level A, the priority training group, after the re-ranking. A further 15 moved from Level A to either Level B or C. The average nominee shifted by just over 20 points in the percentile rankings.

Given these substantial changes, we would like to understand what the managers took into account as a result of the Selection Training. Regressing the change in the percentile ranking against the diagnostic scores discussed as part of the Selection Training provides information on this. The first column of Table 6 below shows the results of this regression, aggregating the seven diagnostic scores plus years of schooling into two indices.[5] The “aptitude” index includes years of schooling, literacy, numeracy, processing speed and garments knowledge. The “attitude” index include family support, interest, and confidence.

We see that factories move nominees on the basis of both aptitude and attitude, but that the movement on attitude is much stronger. This may reflect the fact that the managers have better information about aptitude than they do about attitude. In Appendix 1, we repeat this same regression with each of the eight components included individually. Those results show particularly strong effect for confidence, family support, and numeracy, again suggesting that particular attitudes may be less well known to the managers.

[5] We aggregate by standardizing each of the individual diagnostic scores and then summing them.

In sum, the Selection Training did induce managers to change the rankings of nominees. Sometimes this changed the preference order among nominees who were above the training threshold and remained above that threshold. But often this changed the nominee's assignment from training to not training, or the reverse.

**Table 6: Direct Effects of Factory HR Selection Training**

| VARIABLE | (1) Ranking Change | (2) Completed Training | (3) % Days ASV | (4) Completed Training |
|---|---|---|---|---|
| Aptitude Diagnostic Score (standardized, 5 components) | 0.027** (0.009) | | | |
| Attitude Diagnostic Score (standardized, 3 components | 0.079*** (0.016) | | | |
| Log age of Nominee/ Trainee | 0.210 (0.168) | | | |
| Log factory tenure of Nominee / Trainee | -0.037 (0.028) | | | |
| Factory assigned to Selection Training | | 0.067 (0.065) | -0.037 (0.097) | |
| Nominee percentile rank, pre Selection Training | | | | 0.055 (0.134) |
| Change in rank of Nominee after Selection Training | | | | 0.079 (0.120) |
| Observations | 155 | 201 | 151 | 107 |
| R-squared | 0.264 | 0.007 | 0.002 | 0.166 |
| Factory FE | YES | NO | NO | NO |
| Robust standard errors in parentheses, clustered at factory level. *** p<0.01, ** p<0.05, * p<0.1 | | | | |

**The Effect of Selection Training on Trainee Outcomes**

We have seen that the selection training results in significant re-ranking, with both the aptitude and attitude of nominees taken into account in the re-ranking exercise. We are also interested in knowing whether the Selection Training resulted in higher program completion and promotion rates for trainees ultimately selected. We begin by regressing an indicator that the trainee competed the training program against a measure of whether the factory participated in the HR selection training, though as we have noted, the statistical power for this test is lower than we would like. Referring to table 6, we find in Column 2 that the completion rate was indeed higher, by around 6.7 percentage points (against a rate of 75.5 percent in the factories not receiving the selection training), though with standard errors clustered at the factory level, the effect is not statistically significant (p=0.31). Nevertheless, the effect is large and in the direction that suggests the selection training may have been effective in selecting better trainees.

In Column 4, we find a similar effect if we instead use only the sample of factories receiving the selection training, and ask whether a trainee's pre-training or post-training ranking is a better predictor of program completion. We use the same dependent variable, indicating that the trainee completed the program, and regress this against both the trainee's original percentile ranking and the change in the ranking following the Selection Training. We find that the coefficient on the original ranking is slightly smaller than the coefficient on the change in the ranking, though both are highly insignificant, reflecting, perhaps, the under-powered nature of the analysis. Finally, the results in Column 3 of Table 6 show that trainees in factories participating in the Selection Training completed a slightly smaller (and insignificant) percentage of the trial period days working as an Assistant Supervisor.

Thus, while neither the factory-level average nor the post-training ranking effects are statistically significant, both indicate that the Selection Training moved factories in the desired direction, toward selecting trainees who were more likely to complete the program. This analysis suggests that factories adjusted the selection of trainees to take the diagnostic scores into account when they are informed about those scores. Factories adjust their rankings to train individuals with higher aptitude and (especially) higher attitude scores. We do not find significantly higher training completion or promotion rates among factories that completed the selection training, though the small sample size prevents us from drawing a more definitive conclusion on this point.

## Effect of training on skills

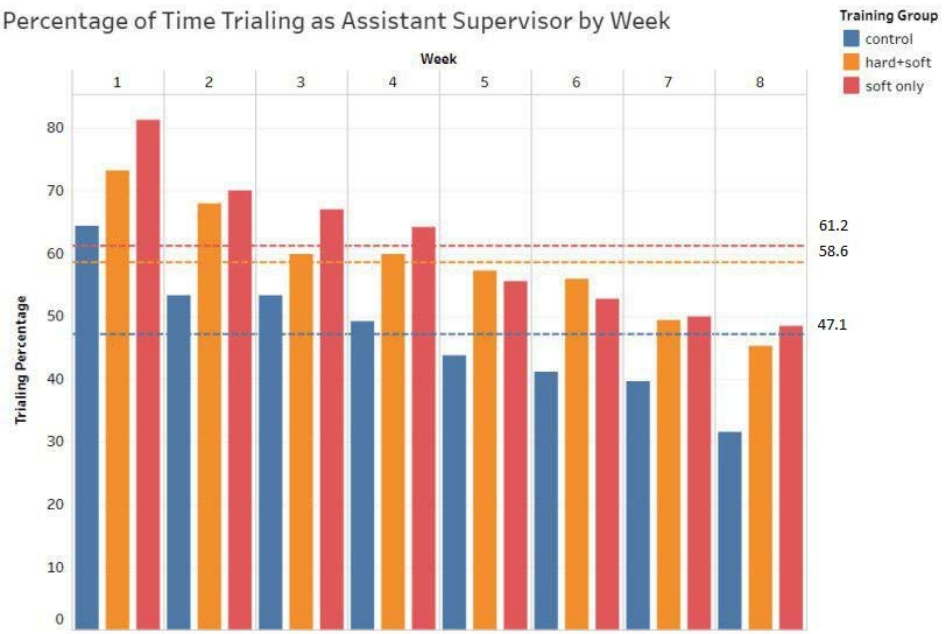
### Technical Knowledge

On average, across trial arms, the trainees completing the program were only trailed as assistant supervisors for about two-thirds of the time they were assigned to, meaning they only received about 5 weeks or 30 days of on-the-job training. This average is lower than we had anticipated. Factory management generally reported that production pressure meant that they needed the trainees, who were generally highly-skilled operators, to work as operators. The implication of this is that many trainees did not get to apply what they learned in the training to the line.

Figure 10 below outlines the percentage of days trialed by week for each treatment group during the 8 weeks that trainees were supposed to trial, including the individuals who dropped out of the program before the end. On average overall, the hard and soft group trialed 55.5% of the time, soft only trialed 61.3% of

the time, and the control trialed 38.4% of the time. During the first week each trial arm had above 60% participation, and the control group drops off faster than the other two groups. However, all three groups see a decline in trialing as an assistant supervisor over the designated period.

**Figure 10: Percentage Time Trialing Weekly**




We tested overall garments knowledge at both baseline and follow-up. Table 7 shows that among trainees there was essentially no change in the measured knowledge. Please we only evaluated here the people who we had scores for both baseline and follow-up. Further, the change in the hard and soft as compared to the soft only and control is not significantly different. This may be partly due to the lack of on-the-job training (trialing), and hence the more limited time to apply the classroom training, as discussed above.

**Table 7: Before and After Garments Knowledge Scores**

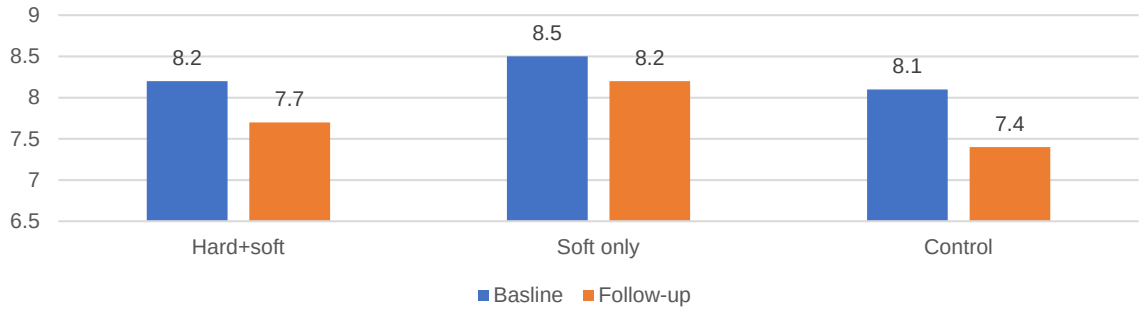
| | Baseline Average Garments Knowledge Score | Follow-up Average Garments Knowledge Score |
|---|---|---|
| **Hard and Soft** | **55.83%** | **55.29%** |
| Observations | 69 | 69 |
| Standard Deviation | 7.32 | 10.06 |
| **Soft Only** | **52.78%** | **53.30%** |
| Observations | 65 | 65 |
| Standard Deviation | 12.21 | 8.95 |
| **Control** | **54.87%** | **54.63%** |
| Observations | 64 | 64 |
| Standard Deviation | 8.49 | 11.05 |

## Confidence in supervisor ability

Previous work training female supervisors has revealed that female operators have lower levels of confidence in the projected ability as supervisors than males being considered for promotion (see Macchiavello et al, 2017.) A six-week training course was sufficient to raise the confidence of the female trainees to a level much closer to that of male trainees. One goal in the design here was to reduce the amount of time required to increase confidence of potential female supervisors.

We measure confidence by first asking trainees how, on a scale of 1 to 10, they would rate their performance as future supervisors for 10 different tasks that are a part of supervisors' jobs – teaching workers new skills, meeting targets, etc. We then ask how they would rate a typical supervisor in their factory on the same tasks. In addition to the 10 ratings on specific tasks, we ask for the overall rating of a typical supervisor and their own performance. Here, we use both the overall rating and the mean of the standardized ratings for the 10 tasks. Figure 11 outlines self-ratings for overall supervisor ability from before and after the trial, excluding dropouts. All groups reduced their self-rating. The control group decreased the most; potentially they had difficulty trialing on the line without the training or lost confidence in their ability after not being trialed on the line as planned.

**Figure 11: Self-rated overall supervisor ability**




When looking at comparisons to typical supervisors, at baseline the trainees rate their own performance slightly higher than a typical supervisor (8.15 vs. 8.01), while at follow-up they rate themselves slightly below (8.01 vs. 8.63). This may reflect the actual experience and challenges of learning the role on the job. We will, of course, be particularly interested in the effect of training on these confidence levels.

To measure the effect of training more closely, we regress the trainee's rating of her own performance at the time of the follow-up survey against her rating for a typical supervisor and her self-rating at the baseline survey. We limit the sample to the 156 trainees who completed training and responded to the questions at both baseline and follow-up. We include in the regression measures of aptitude and attitude (measured at baseline) and the training group to which she was assigned. The coefficient on the training this tell us the number of points on the 10-point scale (or standard deviations for the mean of the 10 tasks) that confidence

changed as a result of participating in either the soft only or hard and soft skills training. The results are shown on Table 8.

We find no significant effect of soft skills only training on confidence (though the measured effect is positive). But we find that the Hard and Soft Skills training increases confidence by 0.65 points on the 10-point scale, or 0.29 standard deviations in the case of the average of 10 skills. The first effect is comparable in magnitude to that reported by Macchiavello et al for the six-week training. We find no effect of the aptitude score in what is effectively a change in confidence (given the inclusion of the baseline measure). And while the baseline attitude measure is strongly correlated with confidence at follow-up, this appears to be driven entirely by the confidence component of this measure, which is effectively another control for baseline confidence levels. We also find no significant effect of working as an ASV for a larger percentage of the trial period days (results not shown on table) on confidence levels.

**Table 8: Confidence in ability as supervisor**

| VARIABLES | (1)<br>Self-ranking, overall supervisor ability | (2)<br>Self-ranking, mean ability 10 tasks |
|---|---|---|
| Baseline own confidence, same measure | 0.256***<br>(0.050) | 0.200*<br>(0.109) |
| Confidence, typical LSVs | 0.343**<br>(0.147) | 0.419***<br>(0.106) |
| Assigned to Soft Skill Training | 0.278<br>(0.257) | 0.103<br>(0.095) |
| Assigned to Soft and Hard Skills Training | 0.651**<br>(0.268) | 0.292*<br>(0.152) |
| Aptitude Diagnostic Score (standardized, 5 components) | -0.041<br>(0.323) | -0.079<br>(0.116) |
| Attitude Diagnostic Score (standardized, 3 components | 0.771***<br>(0.205) | 0.524***<br>(0.116) |
| Log age of Nominee / Trainee | 0.042<br>(1.277) | 0.037<br>(0.407) |
| Observations | 156 | 156 |
| R-squared | 0.505 | 0.543 |
| Factory FE | YES | YES |

Robust standard errors in parentheses, clustered at factory level

*** p<0.01, ** p<0.05, * p<0.1

Notes: The dependent variable in column 1 is the trainees overall rating of her expected performance as a supervisor. The regression controls for the trainee's overall rating of a typical supervisor. The dependent variable in the second column is the mean of the standardized self-ratings in ten specific supervisory tasks; again, the regression controls for both the baseline measure of this same variable and the trainee's average standardized rating of typical supervisors,

# Productivity

We examine three outcomes related to productivity that are reasonably expected to be affected by line supervisors. These are efficiency, alteration rates, and absenteeism. For information on how we calculate these measures, please reference Appendix 4 and for information on terms, please reference our glossary in Appendix 7. We note that the analysis of the production data is particularly preliminary. This reflects the complexity of the data and the time taken to make it uniform across factories and the fact that we are only now completing the collection of the data. It also reflects the need to match the times the trainees were actually working on lines carefully, a task that takes some tome to complete, both because compliance rates were far from perfect and because factories do not record where workers are working each day. We therefore rely on our less frequent phone surveys to construct these measures. We show some initial patterns, but are not yet ready to interpret these in the context of the either the training program or to provide an overall assessment of the effectiveness of the trainees as supervisors. Who expect to provide more information in this regard in the future.

**Efficiency**

Figure 12 shows the trend of the average efficiency across the 60 lines with trainees who were assigned to the hard and soft group. The graph is separated into pre-trial, trial, and post-trial periods. In this case, trial means both when the trainee is trialing on the line. In post-trial, the trainee is either still trialing, is working as an operator, or is working as a full supervisor depending on the factory's decision. However, given the data we use here, the post-trial period covers only the time before the control group receives the classroom

training. So, the differences in the training received by the three groups remain during the post-trial period. Efficiency appears to decrease towards the end of the trial period, but then increases post-trial on the lines where trainees were working.

**Figure 12: Average Efficiency in Hard and Soft across Months**

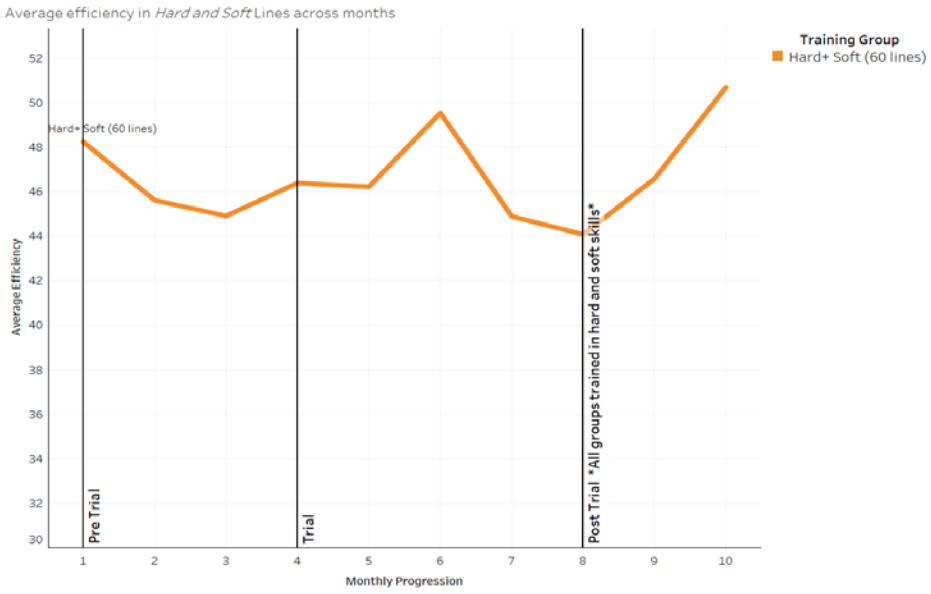
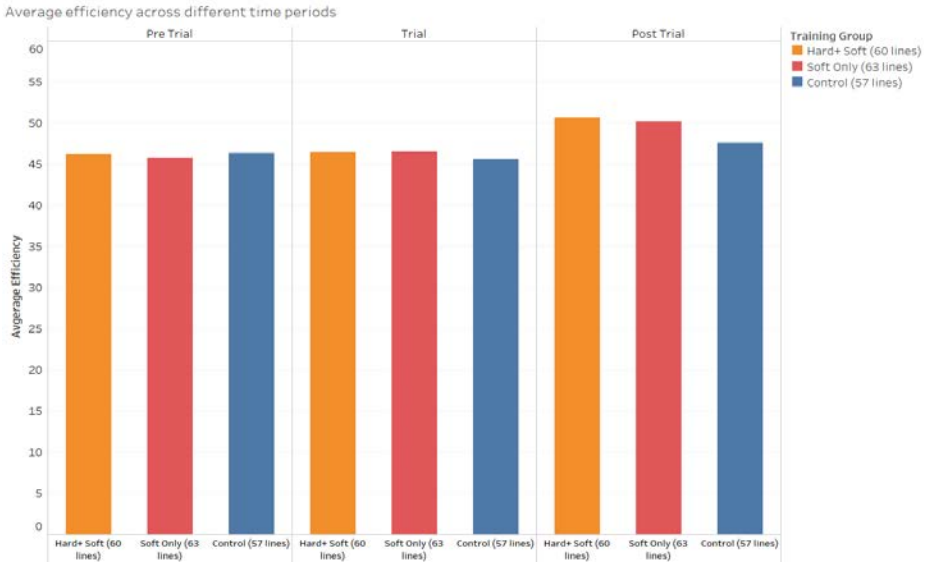



Figure 13 shows the average efficiency for each trial arm in the three phases that we have data for. Efficiency does increase more in the hard and soft as compared to the control, but the difference is not significant.

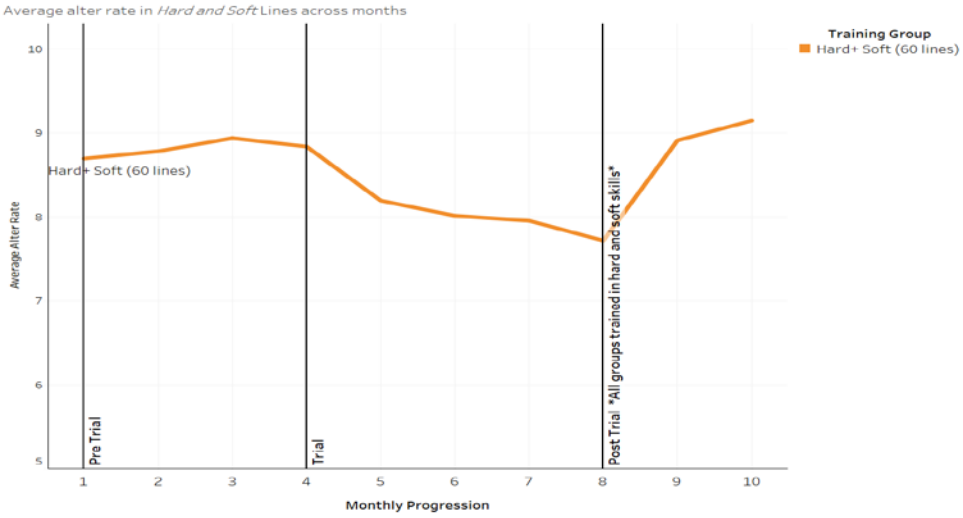
**Figure 13: Average efficiency across periods**




## Alteration Rates

The alteration rate is the number of garments requiring re-work divided by the number of garments checked on the day. In figure 14, improvement is signified by a decline. Across the lines where trainees with hard and soft skills training worked, alteration rates decreased during the trial period and increase afterward.

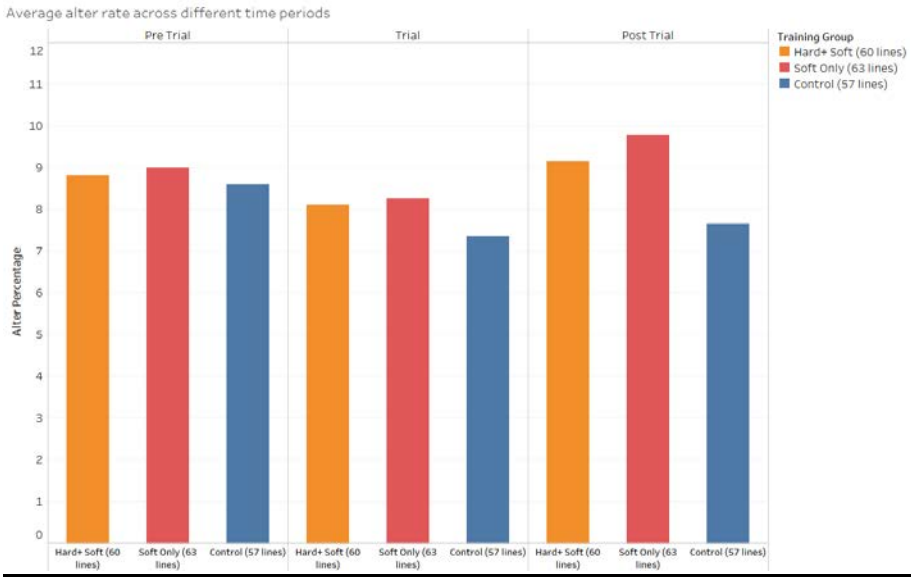
**Figure 14: Average alter rate in hard and soft lines**




As shown in figure 15, during the trial period, the two treatment groups show improvements in alteration rates; however, the control shows a larger improvement. Figure 20 suggests that the training did not

(initially, at least), improve alteration rates. For the two months post-trial, where all groups had received training, there is an increase in alteration rates across the three groups. This may reflect seasonal patterns in the types of products being produced, working hours, or other factors.

**Figure 15: Average alter rate across time periods**




## Absenteeism

Absenteeism is an important measure of supervisor leadership ability and a factor influencing productivity capacity. Among the lines with hard and soft skills trainees, absenteeism increased during the trial period but saw a stark drop post-trial (figure 16).

**Figure 16: Average absenteeism in hard and soft lines**

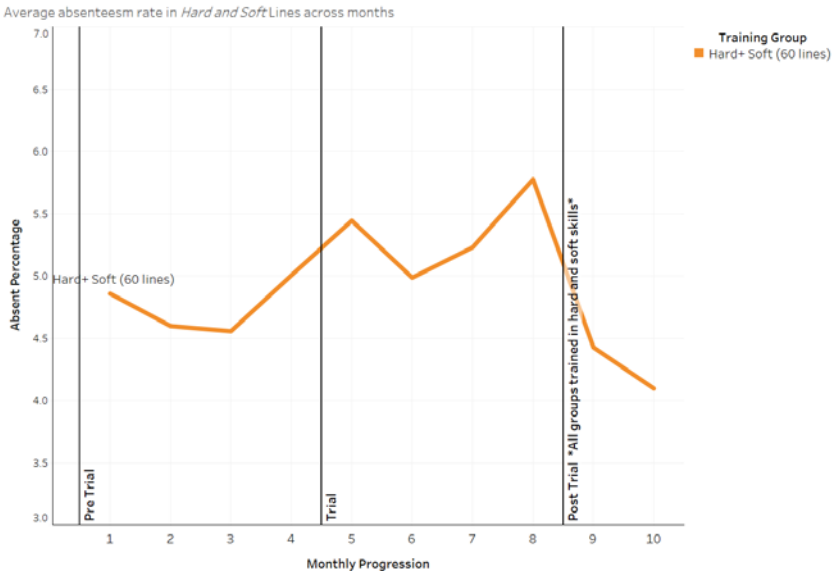
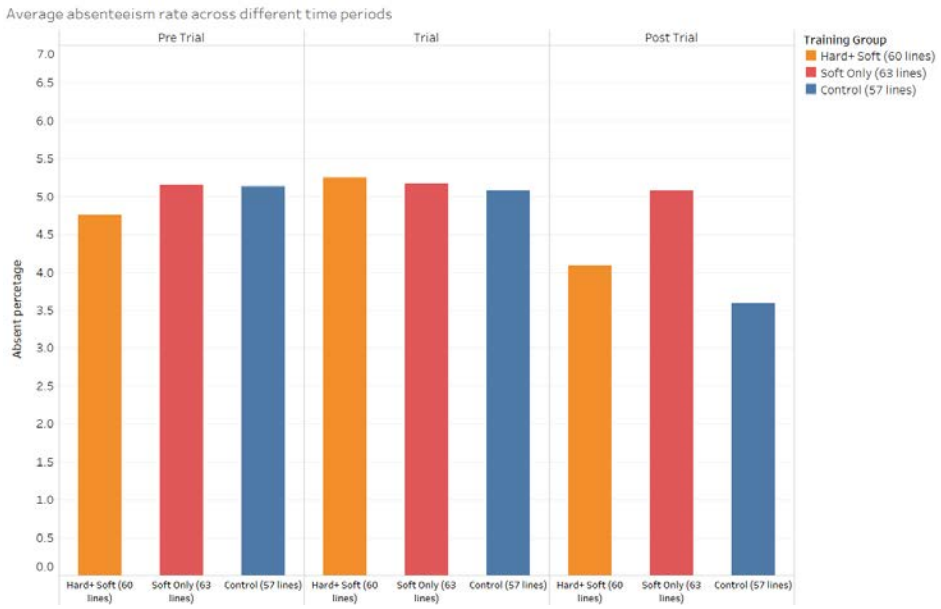


Figure 17 below outlines the average percent absent among the lines by trial arm. While absenteeism among soft skills trainee lines stayed relatively the same, absenteeism decreased among hard and soft skills trainee lines post-trial. However, absenteeism also decreased in the control group.

**Figure 17: Average efficiency across periods**

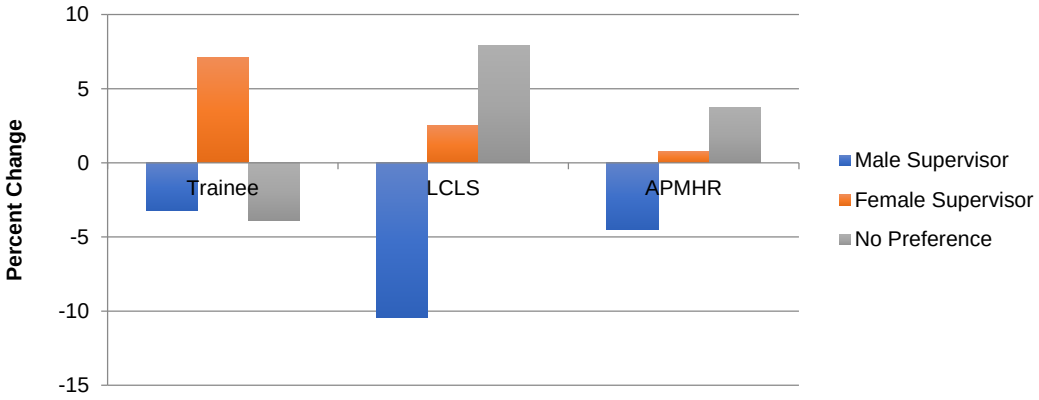



# Perceptions and Support of Trainees

## Perceptions of female supervisors generally

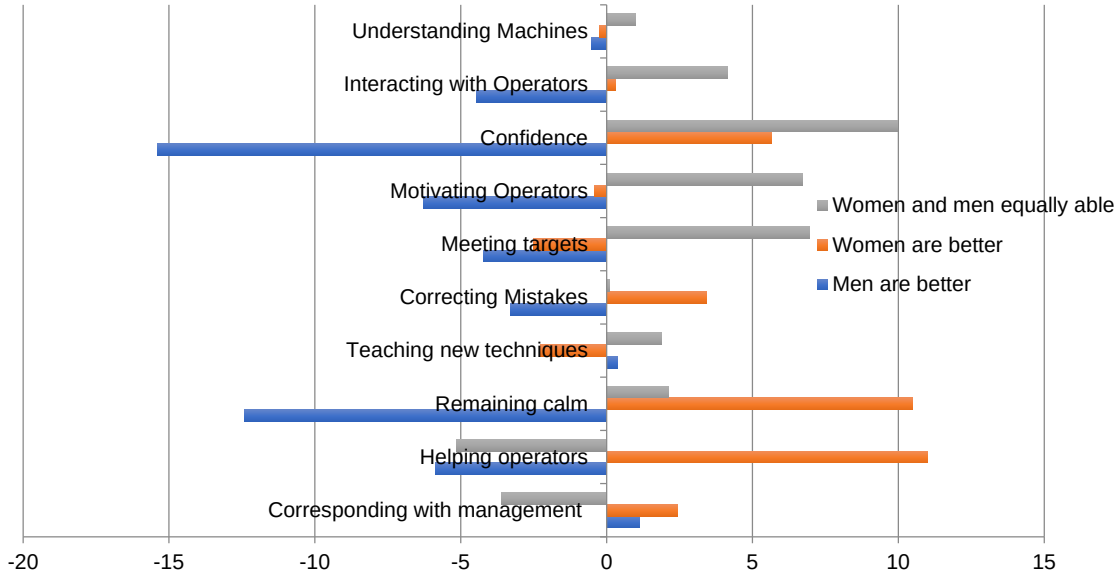
At both baseline and follow-up, we asked workers at all levels of the factory a series of questions related to female and male supervisors generally. One question was “Who do you prefer to work with: male supervisors, female supervisors, or no preference?” Figure 18 shows how the responses to this question changed between baseline and follow-up among trainees, LCLS, and APMHR. Preference for female supervisors increased for all three groups and preference for male supervisors decreased among all three groups. “No preference” decreased among trainees and increased among management personnel (Figure 18).

**Figure 18: Percent Point Difference in Preference from Baseline to Follow-up**




We also asked whether females or males were better at each of 10 specific tasks that supervisors perform as part of their jobs. For each task, respondents were asked whether men were better or women were better in performing the task, or whether there was no difference across gender. With regard to these specific skills, the percentage of respondents saying, “men are better” decreased between baseline and follow-up among LCLS for seven out of ten supervisor skills [see Figure 19 for the full list]. The prevalence of the perception that men are better at “understanding machines,” “corresponding with management” and “teaching new techniques” increased or stayed the same among line managers. Much of this reflects a movement from “men are better” to “both are equal.” But after the training, the percentage of LCLS who reported, “women are better” significantly increased for “confidence,” “remaining calm,” “helping operators,” and “corresponding with management”.

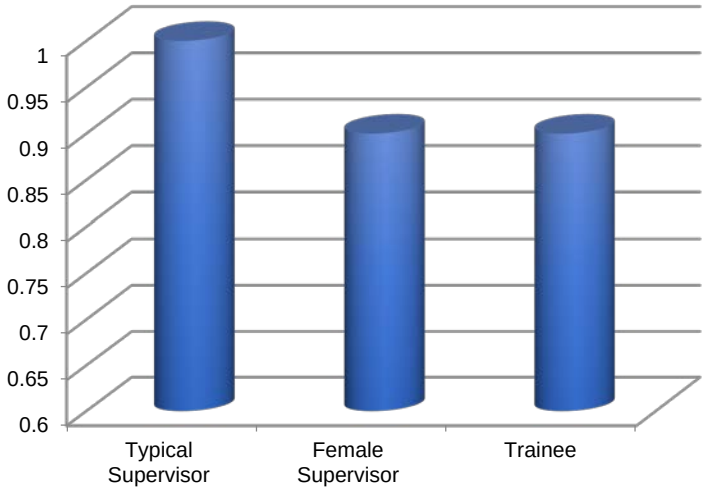
**Figure 19: LCLS Percent Point Change in Perception from Baseline to Follow-up**




## Management Perceptions of Trainees

The surveys also asked managers about abilities of the trainees and of female supervisors more generally. At follow-up, line chiefs and line supervisors (LCLS) reported that their program trainees and a typical female supervisor in their factory both have about 90% of the overall supervisor ability that a typical supervisor has (Figure 20).

**Figure 20: LCLS Perception of overall supervisor ability at follow up**




At follow-up, LCLS felt that the trainees they knew in general improved more quickly than a typical supervisor in their first two months of work (Figure 21). They felt that a female line supervisor generally improves less quickly than a typical new supervisor. Production mangers and HR mangers (APMHR) said that trainees improve at about the same pace as typical new supervisors, and that female line supervisors improve less quickly than typical new supervisors (Figure 22).

**Figure 22: APMHR Response: Improvement in first two months relative to typical supervisor**

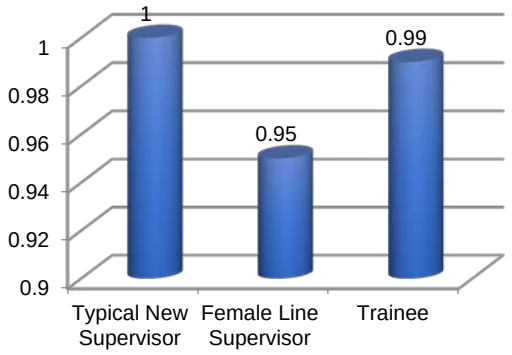
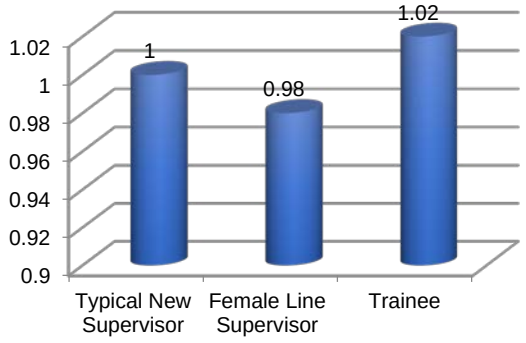



**Figure 21: LCLS Response: Improvement in first two months relative to typical supervisor**




## Factory Support of Trainees

Surveys with operators and management personnel also inquired about perceptions of the level of support the trainees received, particularly around their practicing as an assistant supervisor. Line operators and LCLS responses to questions about factory support for trainees demonstrate a theme of resistance towards the training and towards trainees trailing on the line (figure 23, 24). For instance, over 50% of operators and LCLS both reported that they felt that typical supervisors did not think trainees should be trailing as a supervisor. This is a potential explanation for the low rates of trailing among trainees, as described earlier in the report. About half of operators and LCLS reported other negative attitudes and actions towards trainees: beliefs that trainees are not physically strong enough, less support for trainees than for male counterparts from higher level management, and cases of operator resistance to trainees. About half of respondents shared that operators are generally less responsive to trainee's instructions than to a typical supervisor's instructions and that some trainees had difficulty managing male operators.

**Figure 23: Operator perceptions of resistance from factory**

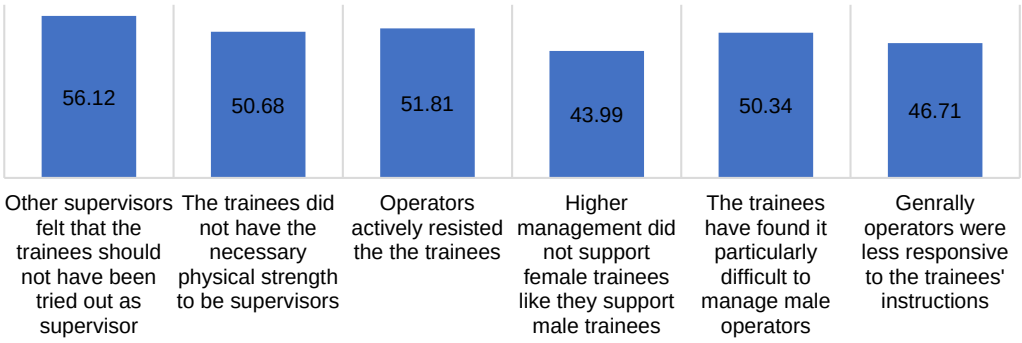
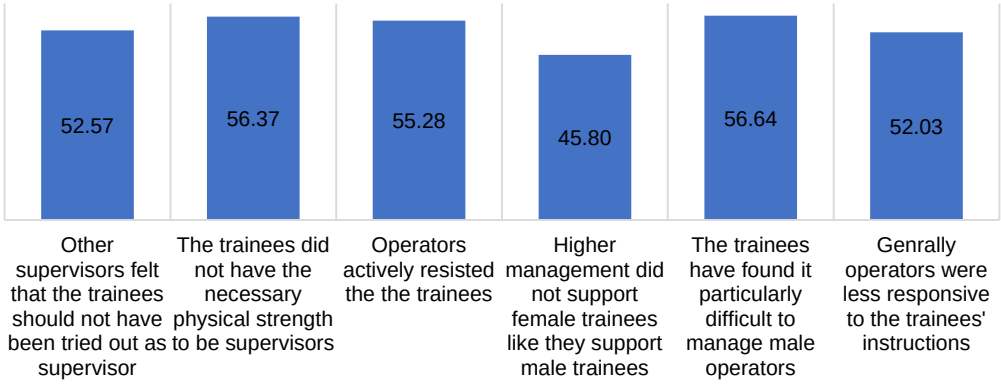
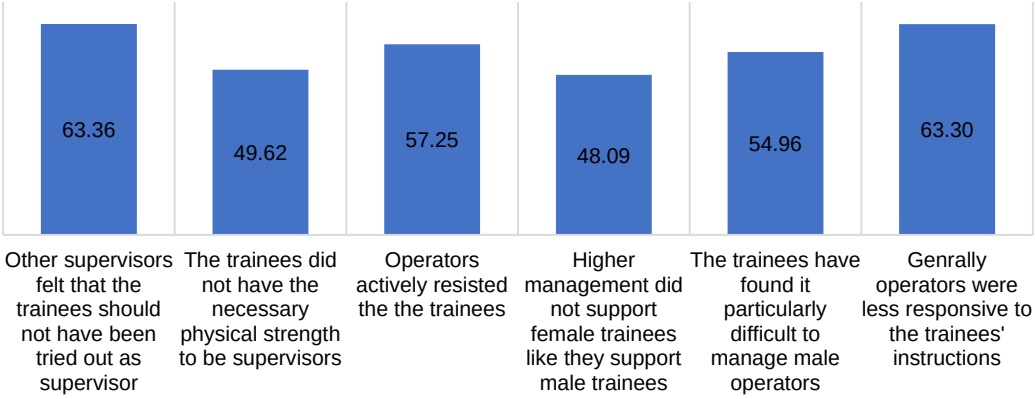


**Figure 24: LCLS Perceptions of resistance from factories**



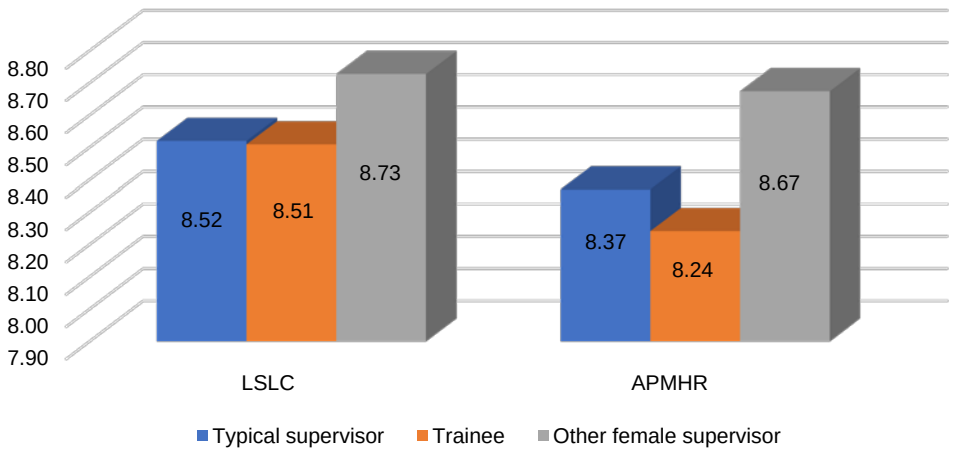

Slightly higher percentages of APMHR than LCLS and operators reported negative actions and attitudes towards trainees (figure 25). Over 60% of APMHR reported that supervisors do not think trainees should be tried out as a supervisor and that operators were not as responsive to trainees as they are to typical supervisors.

**Figure 25: APMHR Perceptions of resistance from factory**



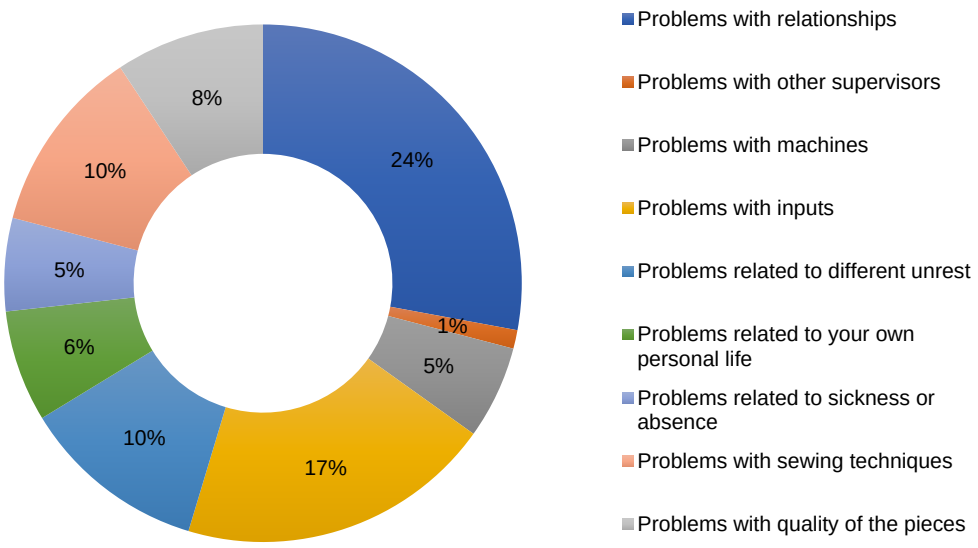

When evaluating overall support, LSLC reported that people supported trainees as much as they support typical supervisors, but not as much as they support female supervisors, as shown in figure 26. APMHR felt that people supported WPT trainees less than they support typical supervisors and female supervisors both.

**Figure 26: Comparison on overall support**




Perhaps in reflection of the above, the most common trainee responses to what the most difficult part of their job was while trialing was problems around relationships with operators and other management personnel (figure 27). This suggests low support for trainees, which could have led to low levels of confidence and interest in promotion.

**Figure 27: Main challenges while trialing as assistant line supervisor**




## Operator Perceptions of Trainees

In addition to management personnel, we also surveyed operators about their perceptions of WPT trainee ability as compared to a typical supervisor and a typical female supervisor. At follow-up, we surveyed four operators on each of the lines where trainees were assigned to work during the trial period. One of our goals was to elicit the operators' views on the performance of the trainees. Similar to the confidence measured we discussed above, we asked operators to rate, on a scale of 1 to 10, the performance of the trainee and of a "typical supervisor in the factory." We asked about the specific tasks that supervisors perform, and for an overall rating.

Operators overall reported that trainees demonstrate similar ability to that of a typical supervisor and are more able than a typical female supervisor (figure 28). Figure 29 breaks down operator perceptions by skill. As with overall ability, trainees are ranked similarly to typical supervisors and better than female supervisors. Operators reported that trainees are better at "remaining calm" than typical supervisors are, and are slightly better at helping and motivating operators.

**Figure 28: Line Operators Overall**

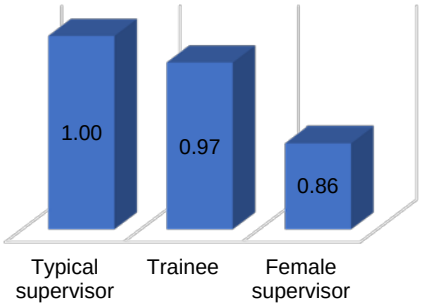
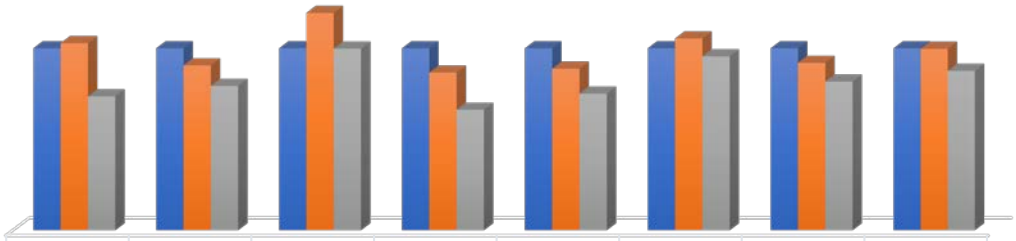



**Figure 29: Line Operator Perception of Trainee Ability**



| | Motivating operators | Being confident as supervisor | Remaining calm | Teaching new techniques | Meeting targets | Helping operators | Correcting mistakes | Corresponding with upper management |
|---|---|---|---|---|---|---|---|---|
| Typical supervisor | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Trainee | 1.01 | 0.97 | 1.07 | 0.95 | 0.96 | 1.02 | 0.97 | 1.00 |
| Female supervisor | 0.91 | 0.93 | 1.00 | 0.89 | 0.92 | 0.98 | 0.94 | 0.96 |

We are interested in understanding whether either the training or the aptitude and attitude of the trainee affect trainee performance as judged by the operators. We are also interested in how characteristics of the operators – for example, whether they are female or male – affects the ratings. Because trainees sometimes worked on different lines, or worked only rarely as an assistant supervisor, we realized that the selected operators may not have been familiar with the trainee assigned to their line. We therefore began by asking them how well they remembered the trainee. For those who did not recalled working with the trainee (around 30 percent of the sample), the questions were instead asked about generic female supervisors. For the

analysis we use only the responses from the 70 percent of the operators who said they remembered working with the trainee. The sample is also limited to those working on lines assigned to trainees who completed the training program.

The first column of Table 9 shows the results of a regression of the operator's rating of the trainee against the operator's rating of a typical supervisor, and several characteristics of the trainee. We standardize both the rating of the trainee and the rating of the typical supervisor, so that the coefficients can be read as magnitudes of effects in standard deviations. We find that trainees with higher scores on the attitude diagnostic are rated significantly higher by operators. Looking at the components of the attitude index, the interest and confidence scores (both measured at baseline) appear to be driving this result. We find no relationship between aptitude and performance as judged by operators. Likewise, we find no effect of either Soft only or Hard and Soft skills training on the subjective performance.

The second column of Table 9 adds two characteristics of the responding operator. The first shows that female operators rate the trainees about one-fifth of a point higher, but the effect is not statistically significant. The second shows that past exposure to female supervisors affects the views of the performance of the trainees. The two-thirds of the operators who report having worked at least two months under a female supervisor previously rate the trainee 0.26 or just over one-quarter of a standard deviation, higher that those without previous sustained experience working under a female supervisor. Previous exposure of less than two months has no significant effect on the ratings (results not shown). The third column adds the percentage of days the trainee worked as an assistant supervisor during the trial period. We find that those working more days are rated more highly by the operators. We are not able to sort out the direction of causation here – it maybe that better trainees worked more often as and ASV, or that those who worked more as an ASV gained skills by doing so. But this at least raises the possibility that compliance with the trialing protocol is important.

Next, we divide the sample into male and female operators to see if they react differently to different characteristics of the trainee. These results are reported in Columns 4 and 5. We find that males are not influenced by past exposure to female supervisors, but are affected more significantly by the attitude – scores for interest and confidence - of the trainee. Female operators, on the other hand, are influenced by past exposure, but not by either the attitude or aptitude of the trainee.

Finally, it is interesting to ask whether the ratings of the operators are correlated with either offers of promotion to trainees, or promotion of trainees. We do this by collapsing the data by taking the averages at the trainee level. As before, we find that being offered a promotion is most strongly associated with the aptitude score, while being promoted is more strongly related to the attitude score. A key right-hand side variable is the rating of the trainee. We take the difference between the standardized rating of the trainee

and the standardized rating of a typical supervisor, and then average this difference across the responding operators. Using this measure, we find that supervisors rated more highly by operators are more likely to be offered promotion (column 6), and more likely to be promoted (column 7). The magnitudes of these effects are significant as well. A one standard deviation increase in the rating of the trainee is associated with a 10-percentage point increase in the likelihood of being offered a promotion and a 17-percentage point increase in the likelihood of being promoted. These results suggest that perhaps either managers take the operator reactions to the trainees into account, or operators and managers both see potential in the same trainees.

**Table 9: Operator ratings of Trainees**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | | | Male operators | Female Operators | | |
| VARIABLES | Op Rating | Op Rating | Op Rating | Op Rating | Offered Promotion | Promoted |
| Aptitude Diagnostic Score (standardized, 5 comp) | -0.113 | -0.111 | -0.144 | -0.065 | 0.253** | 0.152 |
| | (0.098) | (0.100) | (0.169) | (0.128) | (0.126) | (0.119) |
| Attitude Diagnostic Score (standardized, 3 combined) | 0.237*** | 0.227*** | 0.314** | 0.158 | 0.162* | 0.273*** |
| | (0.074) | (0.077) | (0.134) | (0.095) | (0.096) | (0.104) |
| Operator is female | | 0.130 | | | | |
| | | (0.096) | | | | |
| Operator has previously worked >-2 months under a female supervisor | | 0.271** | 0.097 | 0.427*** | | |
| | | (0.102) | (0.162) | (0.128) | | |
| Assigned to Soft Skill Training | -0.069 | -0.054 | -0.294 | 0.143 | 0.024 | 0.102 |
| | (0.129) | (0.121) | (0.222) | (0.145) | (0.159) | (0.168) |
| Assigned to Soft and Hard Skills Training | 0.014 | 0.024 | -0.169 | 0.164 | -0.067 | -0.040 |
| | (0.148) | (0.142) | (0.233) | (0.146) | (0.135) | (0.132) |
| Operator's rating for typical supervisor | 0.272*** | 0.267*** | 0.287*** | 0.274*** | | |
| | (0.051) | (0.053) | (0.082) | (0.064) | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| Log factory tenure of Nominee / Trainee | 0.020 | 0.019 | 0.067 | -0.012 | 0.095 | 0.069 |
| | (0.042) | (0.040) | (0.072) | (0.093) | (0.088) | (0.054) |
| Average dif between rating for trainee & rating for typical SV (standardized) | | | | | 0.104** | 0.179*** |
| | | | | | (0.050) | (0.062) |
| Observations | 513 | 513 | 232 | 281 | 121 | 121 |
| R-squared | 0.189 | 0.208 | 0.260 | 0.254 | | |
| Factory FE | YES | YES | YES | YES | YES | YES |
| Robust standard errors in parentheses ; *** p<0.01, ** p<0.05, * p<0.1 | | | | | | |

Notes: Dependent variable is the rating by the operator of the overall performance of the trainee, on a scale of 1 to 10. The sample is limited to operators who recall working with the trainee, and to trainees who completed the training program. Columns 5 and 6 are probit regressions using data collapsed at the trainee level by taking the average of all responses related to that trainee. This is most relevant for the relative rating of the trainee and typical supervisor, as described in the text.

We are further interested in understanding whether the trainees have a different supervisory style than typical supervisors, and whether training affects their supervisory style. We asked operators whether trainees carried out each of six actions more or less than a typical supervisor. Responses were given on a five-point scale, from “much more frequently” to “much less frequently.” The six actions were: 1) giving more support to less-skilled operators; 2) using praise to motivate operators; 3) using shouting or abusive language to motive operators; 4) solving problems herself rather than consulting superiors; 5) practically demonstrating techniques; and 6) sitting down and sewing when operators are on a toilet break. We standardize and average the responses to these questions as an index of cooperative supervisory style. In doing this, we reverse the scale the scale for all except item 3 (shouting) so that a higher score indicates a more cooperative management style for each item. We also exclude item 4 as it relates to communication with superiors rather than interactions with operators.

**Figure 30: Percent of Operators Report Trainees Do Behavior More Often than Typical Supervisor**

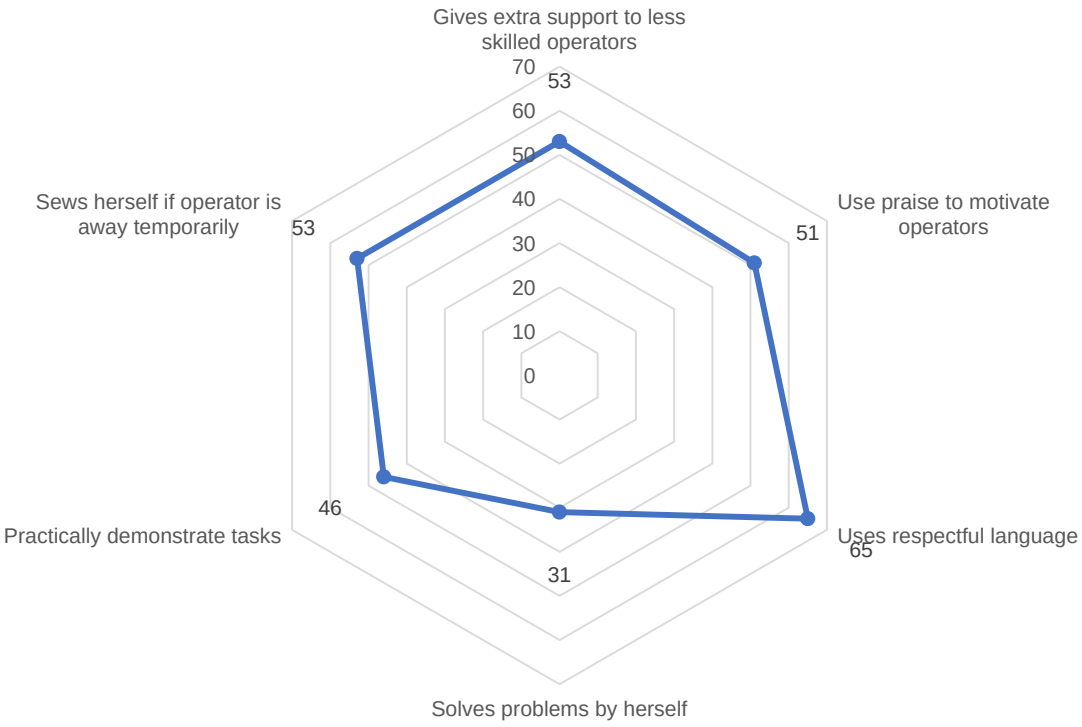



Figure 30 shows that operators report trainees have a more cooperative supervisory style than typical supervisors. Just over half of operators felt that trainees give extra support, motivate operators, and use respectful language more than typical supervisors do. In an unreported regression, we find that trainees with higher attitude diagnostic scores are reported to use the cooperative supervisory style more frequently; the training has no effect on reported supervisory style.

In Table 10, we ask whether and how the cooperative supervisory style is related to the operator rating of the trainee's ability as a supervisor. We find a very large and highly significant relationship that indicates the supervisors who employ the more cooperative supervisory style are also viewed as being more effective supervisors.

**Table 10: Effect of Management Style on Rating of Supervisor**

| | (1) | (2) | (3) |
|---|---|---|---|

| VARIABLES | Operator Rating | Females: Operator Rating | Males: Operator Rating |
|---|---|---|---|
| "Cooperative" supervisory style (standardized) | 0.533*** | 0.537*** | 0.482*** |
| | (0.066) | (0.099) | (0.149) |
| Aptitude Diagnostic Score (standardized, 5 components) | -0.146* | -0.087 | -0.184 |
| | (0.077) | (0.082) | (0.150) |
| Attitude Diagnostic Score (standardized, 3 components | 0.183** | 0.152* | 0.231* |
| | (0.071) | (0.084) | (0.116) |
| Operator is female | 0.155* | | |
| | (0.084) | | |
| Operator has previously worked >-2 months under a female supervisor | 0.245** | 0.357*** | 0.140 |
| | (0.093) | (0.120) | (0.157) |
| Operator's rating for typical supervisor | 0.290*** | 0.297*** | 0.289*** |
| | (0.047) | (0.059) | (0.085) |
| Observations | 521 | 284 | 237 |
| R-squared | 0.318 | 0.372 | 0.314 |
| Factory FE | YES | YES | YES |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Notes: Cooperative management styles is an index of responses to 4 questions asking operators to compare trainees to other supervisors on the amount of help they give to less-skilled operators, the amount of shouting, whether they demonstrate sewing techniques and whether they sew to cover for operators taking toilet breaks. The operator rating is as defined earlier.

In sum, trainees are reported to have a more cooperative supervisory style and more cooperative style is seen to make them more effective supervisors. Selection rather than training appears to be the most important driver of these relationships.

# Conclusions

This report provides insight into the mechanisms around trainee program completion and promotion, challenges to program success, perceptions of females in leadership positions, and the impact of the training itself. The following points summarize our main findings:

- Trainee selection matters. The diagnostic measures are predictive of all of the important outcomes. In particular, the attitude of the trainee at the beginning of the program is strongly associated with completion of the program and holding a supervisory position in the most recent data collected. Attitude is also predictive of the trainee's performance rating as given by operators working under her supervision, and of a more cooperative supervisory style. Family support is particularly important for completion of the training program, but interest in the position and confidence in ability are the more important predictors for other outcomes.
- Training with mid- and high-level managers leads them to choose trainees with higher attitude scores, Given the effect of baseline attitudes on many outcomes, we believe this suggests that the half-day selection training is effective.
- Soft skills training alone is not sufficient to increase the confidence of trainees, nor to affect other outcomes. Soft and hard skills training combined has a significant effect on trainee confidence increasing confidence by 0.65 points on the 10-point scale, or 0.29 standard deviations in the case of the average of 10 skills. However, neither training significantly increased the level of technical knowledge about garment production. Training had only weak (and generally statistically insignificant effects) on other outcomes.
- Overall, our results show that trainees with higher scores on the attitude diagnostic are rated as significantly more effective by operators. Further, we find that trainees rated more highly by operators are more likely to be offered promotion and more likely to be promoted.
- Trainees have a different supervisory style than typical supervisors, one that we term more cooperative. However, there is only very weak evidence that this style comes form the training. Rather, it seems to be embedded in the individuals.
- The program resulted in offers of promotions across factories. Within a cohort of 134 trainees who completed the program and 28 who dropped out, factories made 92 promotion offers as of January 2018, meaning that about half of trainees have been offered a promotion. Out of the 134 program completers who were surveyed in January, 74 (or 55%) are currently working as a supervisor or assistant supervisor. Of 10 factories without any female supervisors at the start of the program, eight promoted at least one trainee during the program.
- In spite of these successes, dropout rates were higher that expected. Moreover, a surprisingly high percentage of those offered promotions to supervisor declined them. Those who dropped out of the program tended to be less confident and less educated with less family support. When looking at self-reported reasons for dropping out and for not accepting offers from promotion, the majority of responses centered around not being interested in the content and responsibilities of the supervisor role.

- Efficiency on the lines where hard and soft trainees worked saw a slightly positive increase, though this was not statistically significant.
- Perceptions of trainees among LCLS, upper-management, and operators were relatively positive. LCLS rated trainee ability similar to a female supervisor and slightly less than a typical supervisor, while operators felt that trainees were about as able as typical supervisors and significantly better than a female supervisor.
- While there are positive perceptions of trainee ability, there is also evidence of resistance towards trainees working as assistant supervisors. About half of LCLS, operators and APMHR reported lack of support for trainees and negative attitudes towards their success as supervisors. Upper management reported that overall people at different levels support WPT trainees less than they support typical supervisors. This could be an explanation for the high dropouts, low acceptances of promotions, and low compliance with trialing on the line.

A potential explanation for the lack of change in efficiency and overall garments knowledge is that it takes time for trainees to absorb and apply skills and secondly that trainees did not have trial on the line the amount of time that was requested and recommended. In past projects, we have seen that it takes several months for new supervisors to catch up to with other lines' productivity. When we compare trainee's lines to other lines, we are comparing trainees to skilled supervisors with experience, so the comparison will not make the trainee lines look as productive.

When evaluating confidence and other measures, it is important to keep in mind that the training is keeping people engaged and working as assistant supervisors. One thing that the training does is get buy-in from both the workers and the factories. This is also reflected in our findings that completing the program is predicted by simply being in the hard and soft training. This "buy-in" will keep in relatively weaker supervisors with lower confidence because they see they are being invested in, and therefore do not pull out of the program. The training may have helped, but the training and the trialing may have decreased some trainee's confidence. Further, those with low confidence at the start likely dropped out early or were eliminated at selection, so it may look like confidence went down in the treatment when maybe it brought it up but we lost people who were not confident to begin with.

The success of the training is largely dependent on enforcement of trialing, selecting participants with positive attitudes, and on working with management to both effectively select trainees and to support them through the training process. Our results imply that female promotion may be more limited by social norms and negative attitudes towards women in supervisor roles (perpetrated by operators, management and families, as well as internalized by women themselves) than technical capacity of female garment workers. However, of the 10 factories who had 0 female supervisors at the start of the program, 9 factories offered promotions to trainees and 8 had at least one trainee accept a promotion. We see that the training contests

these norms to some extent through having an impact on perceptions of female supervisors and through opening up space for offers of promotion to women.

Based on the above, we present the following recommendations:

1) When selecting training participants, management should select nominees who express a strong interest in the position, and have higher levels of confidence and strong family support for becoming a supervisor. Higher years of schooling is also predictive of completing the program.
2) To ensure higher program impact on skills, it is important to ensure that trainees get to on-the-job practice through trialing as assistant supervisors. In this program, trainees only trailed about 60% of the time they were supposed to, which was a large limitation on their ability to make impact on productivity.
3) Soft skills training alone appears not to be effective. The training in soft and hard skills combined is effective in increasing the confidence level of trainees. We cannot say for certain whether the larger effect of the combined training results from the diversity of skills being taught or simply the increased training time. But the evidence supports a longer training period, and the combination of soft and hard skills has important impacts.
4) Finally, our results demonstrate that norms characterizing females as less able to be supervisors remain prevalent among different levels of factory workers. These norms likely discourage female willingness to be promoted. The selection training offered to middle- and upper-level managers included a discussion of norms and implicit biases. The training was effective in getting managers to select trainees with characteristics associated with higher levels of promotion and effectiveness as a supervisor.

# Appendix

## Appendix 1: Diagnostic Scoring of Nominees

A number of different tools were used to assess the abilities of female operators who had been nominated for the Work-Progression & Productivity Toolkit. After the factory has nominated workers, but before the on-boarding session happened, the IPA team visited the factory to assess the abilities of the nominees.

Nominees were only disqualified based on numeracy and literacy scores, but we collected much more data, so that we could examine which nominees succeeded towards the end of the program. The areas that were tested and their respective tests are as follows:

1. Literacy: Multiple choice questions testing reading comprehension, vocabulary, basic grammar, paragraph/letter structure, and writing. The maximum score is 20 points.

2. Numeracy: Multiple choice questions testing calculations, fractions, percentages, number patterns, angles, and visual patterns. The maximum score is 20 points.

3. Processing speed (coding): This test is modeled after the Wechsler Adult Intelligence Scale is called 'Digit Symbol' (WAIS-R), 'Digit-Symbol-Coding' (WAIS-III), 'Coding' (WAIS-IV), also known as the Digit symbol substitution test. It is a neuropsychological test sensitive to brain damage, dementia, age and depression. It consists of digit-symbol pairs (e.g. 1/-,2/┴ ... 7/Λ,8/X,9/=) followed by a list of digits. Under each digit the subject should write down the corresponding symbol as fast as possible. The number of correct symbols within the allowed time is measured. The maximum score is 133 points.

4. Processing speed (Symbol Search): Symbol Search is a subtest of the Wechsler Adult Intelligence Scale (WAIS). The Symbol Search subtest is designed to assess information processing speed and visual perception. High scores require rapid and accurate processing of nonverbal visual information. During Symbol Search, the examinee is asked to mark either the yes or no checkbox with a pencil in response to as many items as possible within 2 min. The maximum score is 60 points.

5. Garments Knowledge: Consists of q21 - q275 in the survey instrument, and contains multiple choice, and open-ended questions testing what machine to use for what operation, names of processes, cause of mechanical issues, cause of quality issues, identifying quality issues in photographs, identifying working condition issues in photographs, and understanding an operation breakdown. The maximum score is 84 points.

6. Family Support: Consists of q811 - q3121_6 in the survey instrument, and gives 5 statements about family support, and asks the respondent to respond on a four-point scale from agree to disagree. An additional three questions ask about the level of support given to other women in the family who work in garment factories. The answers to all questions are recoded so that higher numbers represent more support, and are summed to give a maximum score of 24 points.

7. Interest: The survey instrument has 2 questions about whether they would want to be promoted to supervisor or line chief. Then, we have four questions that indirectly asks whether they are interested in the supervisor position, and asks the respondent to respond on a four-point scale from agree to disagree. The answers to all questions are recoded so that higher numbers represent more interest, and are summed to give a maximum score of 18 points.

8. Confidence: Consists of asking how they would rate their performance compared to a typical supervisor on a 5-point scale. Then, we have three questions that indirectly ask whether the respondent is confident. The respondent is asked to choose between two statements. One statement that says "I am confident" using various words, and a dummy statement about the factory. The answers are recoded so that higher numbers represent more confidence and are summed to give a maximum score of 7 points.

Disqualification rule: No nominee should be disqualified on any category other than literacy and/ or numeracy. The rule is:

- If nominee gets 24% on literacy and 24% on numeracy, she fails
- If nominee gets 24% on literacy and 25% on numeracy, she passes
- If nominee gets 0% on literacy and 100% on numeracy, she fails

| Testing area | Total score |
|---|---|
| Literacy | 20 |
| Numeracy | 20 |
| Processing speed (coding) | 133 |
| Processing speed (symbol search) | 60 |
| Garment knowledge | 84 |
| Family support | 24 |
| Interest | 18 |
| Confidence | 7 |

# Appendix 2: Ranking, Completion, Days Trialed, and Promotion Outcomes with Individual Diagnostic Scores

The below table examines whether the trainee completed the training program, or dropped out before the end based on the eight key measures as described in Appendix 1. The sample excludes trainees in the three factories that dropped out altogether, as we consider those to be factory-level decisions rather than trainee decisions. The regressions control for factory fixed effects, appropriate as the factory was the unit of randomization for the training treatments. The table shows that the family support score is most strongly associated with completion of the training program. Interest is significantly associated with being trialed on the line, while processing speed is associated with being offered a promotion.

Appendix Table 1

| VARIABLES | (1) OLS Ranking Change | (2) Probit Completed Training | (3) OLS % Days ASV | (4) Probit Offered Promotion | (5) Probit Promoted |
|---|---|---|---|---|---|
| Years of Schooling | -0.010 | 0.064** | 0.055** | 0.064 | 0.052 |
| | (0.028) | (0.032) | (0.023) | (0.044) | (0.041) |
| Assigned to Soft Skill Training | | 0.075 | 0.178** | 0.080 | 0.134 |
| | | (0.078) | (0.081) | (0.136) | (0.138) |
| Assigned to Soft and Hard Skills Training | | 0.112* | 0.184** | -0.076 | 0.001 |
| | | (0.064) | (0.073) | (0.133) | (0.137) |
| Literacy score (standardized) | 0.083 | -0.009 | 0.044 | -0.140 | -0.085 |
| | (0.052) | (0.060) | (0.063) | (0.108) | (0.085) |
| Numeracy score (standardized) | 0.085** | 0.002 | -0.063 | 0.083 | 0.088 |
| | (0.030) | (0.060) | (0.042) | (0.118) | (0.080) |
| Processing Speed score (standardized) | -0.059 | 0.006 | -0.047 | 0.221*** | 0.032 |
| | (0.036) | (0.066) | (0.042) | (0.054) | (0.064) |
| Garments Knowledge score (standardized) | 0.062 | -0.012 | -0.015 | 0.066 | 0.021 |
| | (0.042) | (0.027) | (0.033) | (0.072) | (0.062) |
| Family Support score (standardized) | 0.066* | 0.069** | 0.029 | -0.006 | 0.066 |
| | (0.033) | (0.033) | (0.032) | (0.078) | (0.076) |
| Interest score (standardized) | 0.040 | -0.002 | 0.102** | 0.062 | 0.090 |
| | (0.037) | (0.040) | (0.041) | (0.091) | (0.111) |
| Confidence score (standardized) | 0.133*** | 0.042 | 0.016 | -0.040 | 0.034 |
| | (0.031) | (0.042) | (0.032) | (0.066) | (0.073) |
| Observations | 155 | 174 | 184 | 139 | 134 |
| R-squared | 0.351 | | 0.465 | | |
| Factory FE | YES | YES | YES | YES | YES |

Robust standard errors in parentheses, clustered at factory level
*** p<0.01, ** p<0.05, * p<0.1

# Appendix 3: Productivity Data Protection Protocols

| Data Type | Reports |
|---|---|
| Daily Data | Efficiency/Production report<br>Endline quality report<br>Daily input report<br>Daily target report<br>Daily man machine report |

Factory data is used to measure efficiency, absenteeism, and defect rates across lines across factories. The factory data is collected in the shape

| | |
|---|---|
| | Daily attendance<br>Daily absence<br>Individual job card (in-out report) |
| Monthly Data | Salary Sheets<br>Production Planning Documents |
| Static Data | Operation Breakdowns<br>Administrative Data |

that factories record it, so we do not ask factories to do anything special to the data on our behalf. Our final round of data collection is underway now. To date, we have production data from all 27 factories, attendance data from 22 factories, and quality data from all 27 factories. These current set of data spans from pre-intervention to October 2017.

The exact names of the sheets provided vary somewhat across factories. This list is prepared based on our experience of working with RMG factories. Usually, these reports are for the collection of the daily line-wise variables mentioned with the report names.

Confidentially agreements and data handling procedures guarantee the confidentiality of individual employee and factory data, particularly those of study participants. The administrative records are highly confidential and we developed protocols to protect the confidentiality of the data, including encryption and storage in locked safes. To acquire the administrative data, two data collection managers visited each factory together or individually to collect data from the relevant departments. The data were collected in soft copy (through USB or email) or through hard copy duplicates of factory records. In order to minimize trips to the factory, data collection managers encouraged factories to share soft copy data through a Dropbox folder using Boxcryptor. However, some factories had larger amounts of data in hard copy or were not comfortable with online collection methods.

The soft copy data was transferred to encrypted files on a separate secure hard drive that is kept in a locked cabinet outside of working hours. The data that came from a USB were moved immediately after collection at the factory to an encrypted file on the laptop hard-drive, and the data on the USB was deleted. Upon returning to the office, all data collected on laptops were moved to the encrypted factory data hard-drive, and data on laptops were deleted. In the case that data came from emails, the email attachment was saved to an encrypted container on the laptop hard drive and moved to the factory data hard drive as soon as was feasible, after which all attachments saved in laptops were deleted. Data managers confirmed that all email recipients had deleted emails, and that downloaded folders did not contain any data with personally identifiable information.

In the case of hard copies, the data were recorded into an excel template using double entry method and the copies were stored in a locked safe. During the data entry process, the data entry operators anonymized the documents by covering any factory names with a black marker. The electronic files given to the data team had the line names coded. Once the data was entered, the hard copy documents were again filed.

Once data cleaning was complete, the hard copies were moved to a storage facility where they would be stored for three years after project completion.

# Appendix 4: Productivity Data Calculations

For our calculation of productivity, our main indicator we use is line efficiency. Line level efficiency is calculated in a manner consistent with standard practices in the industry. The measure is (sewing output minutes) / (available worker minutes). The sewing output minutes is obtained by multiplying the number of pieces produced on a given day by the standard minute value (SMV) of the garment being produced. The available minutes is the number of production workers (operators plus helpers) assigned to the line multiplied by the number of minutes that the line operated in a day. A typical Bangladeshi factory selling directly to large brands has an average efficiency of 40 to 50 percent. Note that factories may adjust their standard minute value (SMV) measures to the factory context. As a result, the efficiency data may not be directly comparable across factories. We have found that the raw data will reflect the relative efficiency of a factory for most, but not all factories. Adjusting the data to a comparable base takes considerable time and effort.

We also use alteration rates and absenteeism as indicators of productivity changes. The alteration rate is the number of garments requiring re-work divided by the number of garments checked on the day. Absenteeism is an important measure of supervisor leadership ability and a factor influencing productivity capacity, and is just the number or percentage of workers who are absent from a line over a given period of time.

Productivity is also a function of human capital. In the context of this project, the most relevant human capital is the skills of the sewing machine operators and the skills of the line supervisors. We want to understand the extent to which the training that we provided increased the level of human capital of line supervisors. We assess this by measuring the productivity before and after the training across the factory production lines. To evaluate the impact of the training, we will calculate the difference in productivity between the pre-training period and the post-training period between treatment lines and the control lines. Because the trainees assigned to given production lines were randomly allocated to one of the two treatment groups or the control group, this provides a clean measure of the effect of immediate training. However, we are also interested in the productivity of females as a whole. To provide an indication of this effect, we compare the efficiency of lines on which trainees in each of the three groups work with efficiency of all other production lines in the factory.

# Appendix 5: Survey Sample, Tools, and Protocols

Baseline surveys were conducted in November 2016 immediately before training, and the follow-up surveys occurred after the training was completed and before the control was trained.

There are two broad aims of collecting survey data:

1. Assessing the impact of the training on productivity in the factories: While we aim to detect productivity changes that occur in the factories through analysis of factory production data, we also aim to understand the channels through which training affects productivity through data collected in surveys of trainees, managers, and line-level operators. We collect information through surveying on the following dimensions:
    a. Work practices within the factories
    b. Worker attitudes to work and to each other
    c. Worker and manager assessments of trainees, and trainees' self-assessment
    d. Worker relationships with supervisors (including female supervisors) and management more generally

2. Understanding the RMG industry: The data help to provide a broader picture of practices and trends in the sector, including:
    a. Increased knowledge about the type of people that are working in the industry, how much experience they have, their education and family background, as well as their career history in the sector
    b. Improved understanding of existing managerial practices and attitudes toward those practices
    c. Improved understanding of communication within the factory

The survey team selected workers of the following types. All except for nominees were surveyed at both baseline and follow-up (nominees were only surveyed at baseline):

1. Trainees: operators nominated for and selected to participate in the program.
2. Nominees: operators nominated for the program, but who did not participate in the program either because they did not pass the initial numeracy / literacy screen or because they were not ranked high enough by the factories.
3. Trial line operators: workers operating the sewing machines on one of the lines where trainees were assigned to work or were working. An operator is associated with a single line.
4. Trial line Supervisors: The lowest level of management, their role is to supervise the machine operators and helpers and all aspects of the production process on one of the lines where trainees

were assigned to work or were working. Typically, there are multiple line supervisors on a line, each having responsibility for a part of one line.

5. Trial line Chiefs: The managers immediately above the supervisors in the hierarchy. Their role is to oversee production on the entire line or lines to which they are assigned, and to oversee the work of the supervisors.
6. Production and HR Managers: higher-level positions overseeing line chiefs and supervisors. These managers typically have a large influence on factory promotion policies.

The below table outlines the number of participants surveyed for each of the survey instruments:

| Sample size for both baseline and follow up surveys | |
|---|---|
| Survey Tool | Number of Participants |
| Assistant Production/HR Manager Survey | 156 (120 assistant production manager, 36 HR Manager) |
| Line chief/line supervisor Survey | 447 (246 line supervisors, 201 line chief) |
| Line operator survey (in-person) | 1,023 |
| Line operator survey (by phone) | 511 |
| Trainee Survey | 204 (including those who dropped out; the final enrollment is 144) |
| Weekly Trainee Survey (phone) | 220 (including dropouts and all the replacements) |
| Trainee Diagnostic (phone) | 220 (including dropouts and all the replacements) |

The types of surveys and their respective dates are outlined in table A, while the objectives of each table are outlined in table B.

| Table A: Survey Instruments Dates | | |
|---|---|---|
| Round | Dates | Survey Instruments |
| Baseline | November 2016 - March 2017 | Assistant Production/HR Manager survey<br>Line chief/line supervisor survey<br>Line operator survey (in-person)<br>Line operator survey (phone)<br>Trainee survey |
| Progress surveying | January 2017-<br>August 2017 | Trainee Diagnostic (phone)<br>Trainee weekly survey (phone) |

| | | |
|---|---|---|
| Follow-up | April 2017 – October 2017 | Assistant Production/HR Manager survey<br>Line chief/ line supervisor survey<br>Line operator survey (in-person)<br>Line operator survey (phone)<br>Trainee survey |

| Table B: Survey Instruments Guide | |
|---|---|
| Assistant Production/HR Manager Survey | To measure before and after training APMHR attitudes towards female supervisors, line management, worker motivation strategies and assessments of trainees. The survey also asks about factory characteristics. |
| Line chief/line supervisor Survey | To measure before and after training LCLS attitudes and perceptions towards their operators' work ethic and motivations, supervisor communications, operator trainings and support, and qualities of a strong supervisor. |
| Line operator survey (in-person) | To measure before and after training line operator wellbeing and work satisfaction (non-sensitive questions), as well as attitudes towards female supervisors and assessments of trainees. |
| Line operator survey (phone) | To measure before and after training line operator wellbeing and work satisfaction (sensitive questions). |
| Trainee Survey | To measure before and after training confidence levels of trainees, ability to handle stressful work situations and relations, attitudes towards promotion, women's autonomy around household decision-making, and relationship with management. Additionally, this survey documents trainee anxiety levels. |
| Weekly Trainee Survey (phone) | This short survey asks trainees if they've been offered a promotion, to what position they were offered a promotion, and if they accepted the promotion. It additionally asks if they changed lines or worked in another position. The survey documents challenges in the last week at work. |
| Trainee Diagnostic (phone) | To measure trainee technical knowledge before and after the training as well as confidence and ability to handle stressful work situations and relations. |

All surveying is held to the ethical and confidentiality standard of the Innovations for Poverty Action Institutional Review board. No children (those under the age of 18) were surveyed in these samples.

# Appendix 6: Session Dates

**Hard & Soft**

| Session | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| day_1 | 01_Jan_2017 | 23_Jan_2017 | 19_Feb_2017 | 19_Mar_2017 | 25_Apr_2017 |
| day_2 | 02_Jan_2017 | 24_Jan_2017 | 20_Feb_2017 | 20_Mar_2017 | 26_Apr_2017 |
| day_3 | 08_Jan_2017 | 29_Jan_2017 | 26_Feb_2017 | 27_Mar_2017 | 30_Apr_2017 |
| day_4 | 09_Jan_2017 | 30_Jan_2017 | 27_Feb_2017 | 28_Mar_2017 | 02_May_2017 |
| day_5 | 12_Jan_2017 | 05_Feb_2017 | 05_Mar_2017 | 02_Apr_2017 | 04_May_2017 |

| | | | | | |
|---|---|---|---|---|---|
| day_6 | 16_Jan_2017 | 12_Feb_2017 | 13_Mar_2017 | 09_Apr_2017 | 07_May_2017 |
| day_7 | 17_Jan_2017 | 13_Feb_2017 | 14_Mar_2017 | 10_Apr_2017 | 08_May_2017 |
| day_8 | 31_Jan_2017 | 19_Feb_2017 | 19_Mar_2017 | 16_Apr_2017 | 14_May_2017 |
| day_9 | 01_Feb_2017 | 20_Feb_2017 | 20_Mar_2017 | 17_Apr_2017 | 15_May_2017 |

**Soft Only**

| Session | 1* | 1** | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| day_1 | 28_Dec_2016 | | 24_Jan_2017 | 22_Feb_2017 | 21_Mar_2017 | 23_Apr_2017 |
| day_2 | 29_Dec_2016 | | 25_Jan_2017 | 23_Feb_2017 | 22_Mar_2017 | 24_Apr_2017 |
| day_3 | 03_Jan_2017 | | 31_Jan_2017 | 07_Mar_2017 | 28_Mar_2017 | 02_May_2017 |
| day_4 | 04_Jan_2017 | | 01_Feb_2017 | 08_Mar_2017 | 29_Mar_2017 | 03_May_2017 |
| day_5 | 09_Jul_2017 | 03_Aug_2017 | 09_Jul_2017 | 03_Aug_2017 | 27_Aug_2017 | 08_Oct_2017 |
| day_6 | 10_Jul_2017 | 06_Aug_2017 | 10_Jul_2017 | 06_Aug_2017 | 28_Aug_2017 | 09_Oct_2017 |
| day_7 | 11_Jul_2017 | 07_Aug_2017 | 11_Jul_2017 | 07_Aug_2017 | 29_Aug_2017 | 10_Oct_2017 |
| day_8 | 12_Jul_2017 | 08_Aug_2017 | 12_Jul_2017 | 08_Aug_2017 | 30_Aug_2017 | 11_Oct_2017 |
| day_9 | 13_Jul_2017 | 09_Aug_2017 | 13_Jul_2017 | 09_Aug_2017 | 31_Aug_2017 | 12_Oct_2017 |

**Control**

| Session | 1* | 1** | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| day_1 | 11_Jun_2017 | 10_Aug_2017 | 11_Jun_2017 | 10_Aug_2017 | 20_Aug_2017 | 02_Oct_2017 |
| day_2 | 12_Jun_2017 | 12_Aug_2017 | 12_Jun_2017 | 12_Aug_2017 | 21_Aug_2017 | 03_Oct_2017 |
| day_3 | 13_Jun_2017 | 16_Aug_2017 | 13_Jun_2017 | 16_Aug_2017 | 22_Aug_2017 | 04_Oct_2017 |

| | | | | | | |
|---|---|---|---|---|---|---|
| day_4 | 14_Jun_2017 | 17_Aug_2017 | 14_Jun_2017 | 17_Aug_2017 | 23_Aug_2017 | 05_Oct_2017 |
| day_5 | 18_Jun_2017 | 20_Aug_2017 | 18_Jun_2017 | 20_Aug_2017 | 27_Aug_2017 | 23_Sep_2017 |
| day_6 | 19_Jun_2017 | 21_Aug_2017 | 19_Jun_2017 | 21_Aug_2017 | 28_Aug_2017 | 24_Sep_2017 |
| day_7 | 20_Jun_2017 | 22_Aug_2017 | 20_Jun_2017 | 22_Aug_2017 | 29_Aug_2017 | 25_Sep_2017 |
| day_8 | 21_Jun_2017 | 23_Aug_2017 | 21_Jun_2017 | 23_Aug_2017 | 30_Aug_2017 | 26_Sep_2017 |
| day_9 | 22_Jun_2017 | 24_Aug_2017 | 22_Jun_2017 | 24_Aug_2017 | 31_Aug_2017 | 01_Oct_2017 |

Session 1 soft only trainees were separated into two groups during technical training
* 2 of 6 Session 1 factories merged with session 2 training
** remaining 4 Session 1 factories merged with session 3 training

# Appendix 7: Glossary of Terms

| | |
|---|---|
| Absent | "# of days worker was absent this month" |
| Actual manpower total | "From production report: Actual # helpers and operators on line on the day" |
| Advance | "Deduction for previously given advance (usually before Eid)" |
| Alter | "# of pieces with defects or # of defects (can be >1 per piece)" |
| Annual leave | "Annual leave available" |
| Attendance | "Days attended" |
| Attendance bonus | "Attendance bonus amount" |
| Available minutes | "Actual manpower total * hours * 60min" |
| Basic | "Basic salary" |
| Buyer | "Buyer (coded) of garment produced" |
| Casual leave | "Casual leave (law: 10 days/year with full pay)" |
| Conveyance | "Transport allowance" |
| Day input | "# pieces entering line" |

| | |
|---|---|
| Day output | "# pieces finished on day" |
| Days paid | "Days salary is paid for" |
| Deduction hours | "Hours deducted from workers' working hours" |
| Deduction hours amount | "Deducted amount for being penalized with deducted hours" |
| Department | "Department to which worker is assigned" |
| Designation | "Job position of worker" |
| Dob | "Date of birth" |
| Doj | "Date of joining factory" |
| Dor | "Date of resignation from factory" |
| Earned leave | "Earned/annual leave (law: 1day per 18d of work in previous year if tenure>1yr)" |
| Extra overtime amount | "Extra overtime amount paid" |
| Extra overtime hours | "Extra overtime hours" |
| Extra overtime rate | "Extra overtime pay rate" |
| Fact code | "Factory code (uniquely identifies factory in RMGPP)" |
| Festival leave | "Festival holidays (law: 11 days/year)" |
| Floor | "Floor where worker is assigned" |
| Food allowance | "Food allowance" |
| Gender | "1:male, 2:female, worker gender" |
| Grade | "Worker grade" |
| Gross pay | "Gross pay (usually = basic salary/pay + allowances, but not always)" |
| Hl absent | "From HR data: # helpers assigned to line absent on the day" |
| Hl actual | "From production report: Actual # helpers on line on the day" |
| Hl present | "From HR data: # helpers assigned to line present on the day" |
| Hl registered | "From HR data: # helpers assigned to line" |
| Holidays | "# of holidays this month (weekly off-day, other holidays)" |
| Hours | "Line operating hours" |
| House allowance | "Housing allowance" |
| Id | "Worker id/card number in factory" |
| Incentive allowance | "Bonus for meeting production target" |
| Input date | "Date inputs were provided to line" |

| | |
|---|---|
| Iron absent | "From HR data: # iron workers assigned to line absent on the day" |
| Iron present | "From HR data: # iron workers assigned to line present on the day" |
| Iron registered | "From HR data: # iron workers assigned to line" |
| item | "Type of garment" |
| Late days | "# of late days" |
| line | "Line name" |
| Line code | "line code (fact_code + line_serial) (uniquely identifies factory in RMGPP)" |
| Lo absent | "From HR data: # operators assigned to line absent on the day" |
| Lo actual | "From production report: Actual # operators on line on the day" |
| Lo present | "From HR data: # operators assigned to line present on the day" |
| Lo registered | "From HR data: # operators assigned to line" |
| Loan deduction | "deduction for repayment of loan given by factory" |
| Lost mins | "# Minutes line was down (e.g., power cut, input delay, broken machine)" |
| Lunch out days | "# Days worker did not get back to factory after lunch time" |
| Lunch out deduction | "Deduction for not returning to the factory after lunch" |
| machines | "# Machines on line" |
| Maternity leave | "Maternity leave (law: 16 weeks if tenure > 6 months)" |
| Medical allowance | "Medical allowance" |
| month | “Month (from date)" |
| name | "Worker name" |
| Net pay | "Net pay (usually gross pay - deductions excluding extra overtime amount but not always)" |
| Night allowance | "Night shift allowance" |
| Night day | "# of night shift days this month" |
| Npt reasons | "Causes of non-productive time in sewing line" |
| Order no | "Order name/number" |
| Order quantity | “# of pieces in the order" |
| Ot amount | "Overtime amount paid" |
| Ot hours | "Overtime hours" |
| Ot rate | "Overtime pay rate" |
| Output minutes | “Daily output* SMV" |
| Process | "Name of sewing process / operation" |

| | |
|---|---|
| Reject | "# of pieces rejected (will not be fixed)" |
| Running days | “# days producing the current style" |
| Section | "Section where worker is assigned" |
| Sick leave | "Sick leave (law: 14 days/year with full pay)" |
| Skill bonus | "Bonus for achieving a certain skill-level" |
| SMV | “Standard minute value reported by factory" |
| Spot | "# of pieces with one or more spots" |
| Stamp | "Deduction for stamping salary paper with official seal" |
| Style | “Factory style code (not cleaned)" |
| Target minutes | "Target * SMV" |
| Tax deduction | "Deduction for income tax" |
| Tenure bonus | "Bonus paid to workers for their service year(s)" |
| Total absent | "From HR data: # helpers and operators assigned to line absent on the day" |
| Total allowance | "All allowances combined" |
| Total daily rate | "Basic daily pay rate" |
| Total deduction | "All deductions combined" |
| Total leave | "Total leave days taken this month" |
| Total present | "From HR data: # helpers and operators assigned to line present on the day" |
| Total target | "Daily target" |
| Total check | "# Pieces for which quality was checked that day" |
| Total qc pass | "# Pieces passed quality check that day" |
| Unit | "Unit/area of factory" |
| Weekend holidays | "# of weekend holidays this month (law: 1 day per week) |
| Without pay deduction | "Amount deducted for being penalized to take leave without pay" |
| Without pay leave | "# Days worker was on leave without getting salary" |
| Year | "Year (from date)" |