WORKING PAPER Nº IDB-WP-1792

Overcoming the Gender Bias in Training:

An Empirical Approach in the Latin American Quick-Service Restaurant Industry

Achyuta Adhvaryu
Anant Nyshadham
Jorge Tamayo
Teresa Molina
Sonia Bhalotra

Inter-American Development Bank
Gender and Diversity Division
Department of Research and Chief Economist

April 2026

Overcoming the Gender Bias in Training:

An Empirical Approach in the Latin American Quick-Service Restaurant Industry

Achyuta Adhvaryu*
Anant Nyshadham**
Jorge Tamayo***
Teresa Molina ****
Sonia Bhalotra*****

* UC San Diego, NBER, J-PAL, Good Business Lab & BREAD
** University of Michigan, NBER, J-PAL, Good Business Lab & BREAD
*** Harvard Business School, Digital Reskilling Lab at Harvard
**** University of Hawaii at Manoa
***** University of Warwick

Inter-American Development Bank
Gender and Diversity Division
Department of Research and Chief Economist

April 2026

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Overcoming the gender bias in training: an empirical approach in the Latin American quick-service restaurant industry / Achyuta Adhvaryu, Anant Nyshadham, Jorge Tamayo, Teresa Molina, Sonia Bhalotra.
p. cm. — (IDB Working Paper Series ; 1792)
Includes bibliographical references.
1. Employee training directors-Colombia. 2. Restaurateurs-Colombia. 3. Labor mobility-Colombia. 4. Gender mainstreaming-Colombia. I. Adhvaryu, Achyuta. II. Nyshadham, Anant. III. Tamayo, Jorge. IV. Bhalotra, Sonia R. V. Inter-American Development Bank. Gender and Diversity Division. VI. Series.
IDB-WP-1792

http://www.iadb.org

Abstract

Using administrative and survey data from a large restaurant chain in Colombia, we documented gender differences in the managerial approaches most strongly associated with performance. Based on these findings, we designed two customized curricula: a relationship-based curriculum highlighting characteristics exemplified by high-performing female managers and a task- and metric-based curriculum emphasizing characteristics exemplified by high-performing male managers. Both curricula also included training on universally effective practices. We then implemented a randomized controlled trial in which stores were assigned to a control group, a relationship curriculum group, or a task and metric curriculum group. Managers assigned to the relationship-focused curriculum demonstrate significantly larger improvements on relationship-related questions than do those assigned to the task-focused curriculum, while the reverse is true for task-related questions. Importantly, these targeted gains do not come at the expense of general learning: managers in both treatment groups also demonstrate improvements in questions covering content common to both curricula. Follow-up data collection will measure treatment effects on store- and worker-level outcomes.

Keywords: Gender, managerial practices, training program, Latin America
JEL: L2, M50, J24, L66

1 Introduction

Men and women behave differently in the workplace, often making different decisions and exhibiting different preferences across a range of contexts (Mas and Pallais, 2017; Goldin, 2015; Haegele, 2021; Goldin and Katz, 2011). These differences are particularly relevant in management roles, where the way individuals approach their work directly shapes both their own performance and that of their subordinates (Bloom et al., 2019; Ranganathan and Shivaram, 2021). A large body of empirical work highlights the critical role of managerial skills and practices—particularly in personnel management—in shaping worker productivity (Bloom and Van Reenen, 2007, 2011; Adhvaryu et al., 2018; McKenzie and Woodruff, 2016; Adhvaryu et al., 2022; Frederiksen et al., 2020; Hoffman and Tadelis, 2021). More recent research has highlighted that the managers themselves—their personal characteristics and attitudes—are also important for firm outcomes (Metcalfe et al., 2023; Friebel et al., 2025; Beam et al., 2025). Given the performance implications of managerial behavior and characteristics, understanding how managerial styles differ by gender—and how these differences may affect the daily work of female managers—is essential. These dynamics may also help explain persistent barriers to women's advancement into leadership positions (Feldberg, 2022).

Despite growing evidence on what constitutes effective management, organizations may not always reward the practices that drive performance—either because these practices are difficult to observe or because their value is not fully recognized. Managerial styles and behaviors vary significantly within firms (Adhvaryu et al., 2023c), and those that contribute most to outcomes are not necessarily the most visible to top leadership. These challenges are even more pronounced when the effectiveness of certain practices depends on the identities of both managers and their teams. Gender, in particular, plays a key role: women remain underrepresented in leadership positions, even in sectors where they constitute the majority of frontline workers. This underrepresentation may limit firms' ability to identify, appreciate, and replicate gender-specific management strategies. Prior research highlights that female managers are more likely to devote time to scheduling and accommodate employee preferences, which in turn fosters stronger team cohesion and significantly higher sales—especially in gender-balanced teams (Adhvaryu et al., 2023a).

Do men and women differ in their managerial styles and practices? Do these differences translate into distinct effects on performance? Should firms seeking to maximize productivity encourage gender-tailored approaches to management? And could training managers in styles less common within their own gender group yield additional gains? Answering these questions requires overcoming substantial empirical challenges. It demands observing male and female managers in the same role, within the same industry, operating under identical conditions—then testing whether systematic differences in behavior and performance emerge.

We address these questions in two stages. First, we take a data-driven approach to examine whether male and female managers differ in their practices and behaviors and whether these differences are associated with variation in productivity within the quick-service restaurant (QSR) industry. Using survey and performance data, we identify the characteristics most strongly associated with higher performance for male and female managers, respectively. Second, we use these insights to design gender-informed training curricula and implement a randomized intervention to test whether aligning training content with gender-specific styles

influences managers' comprehension and retention of key concepts. Each curriculum is tailored to the traits linked to higher productivity by gender and adapted from a previously implemented managerial training program (Adhvaryu et al., 2023b).

Building on the design of the gender-informed curricula, we conduct a cluster randomized controlled trial to evaluate their impact. The trial includes three arms: one in which managers receive a task and metric curriculum (designed based on traits predictive of productivity for male managers), another in which they receive a relationship-focused curriculum (designed based on traits predictive of productivity for female managers), and a control group that receives no training. Randomization is carried out at the store level, stratified by store location and by the gender composition of the store's workforce. The intervention is implemented in partnership with a large QSR firm and rolled out across all stores in Colombia.

While a growing body of literature shows that improving managerial practices can boost productivity, most of this work has focused on sectors and settings outside of Latin America. Notably, Adhvaryu et al. (2023b) demonstrate that a managerial best practices training program for frontline supervisors in South Asia increased line productivity by 5.8 percent six months after implementation, reinforcing earlier findings that link managerial quality to firm performance (Bloom et al., 2013). However, similar approaches have not yet been tested in the Latin America and Caribbean (LAC) region, and little is known about how gender-informed designs might enhance the effectiveness of such programs.

This paper contributes to both of these gaps by implementing a tailored management training intervention in the QSR sector in Colombia. We find that training improves managers' understanding of targeted content, particularly in areas where baseline familiarity was lower. Managers assigned to the task- and metric-focused curriculum showed greater gains in content specific to that track, while those in the relationship-focused curriculum exhibited stronger gains in related content. These learning patterns suggest that tailoring training to distinct management approaches can reinforce underdeveloped skill sets without compromising broader knowledge acquisition: both groups also improved in common content areas. Overall, the findings highlight the value of differentiated training strategies in building managerial capacity and promoting more inclusive leadership development.

To complement our short-term knowledge assessments, we plan to collect administrative data in the months following the training intervention. This follow-up will allow us to assess whether improvements in managerial knowledge and practices translate into meaningful changes in operational and workforce outcomes. Specifically, we will analyze impacts on sales, productivity, employee turnover, absenteeism, and other store-level performance indicators. Additionally, we will track changes in workers' perceptions of their managers using the same dimensions captured in a baseline survey, including rapport, psychological safety, harassment, and job satisfaction. These outcomes will provide a more comprehensive understanding of the effectiveness and durability of the training program.

The rest of the paper is organized as follows. Section 2 describes the institutional context, the structure and content of the training curricula, and the experimental design and randomization strategy. Section 3 outlines the survey data and presents key summary statistics. Section 4 details the empirical strategy and presents the main findings, including overall training effects and topic-specific improvements. Section 5 concludes with a discussion of the

implications for gender-informed training and managerial development.

2 Context, Program, and Experimental Design

The LAC region represents the fourth-largest market in the world for food industry sales, generating over $243 billion in annual revenue for a population exceeding 646 million people (Reinhardt et al., 2020). Our implementing partner is one of Latin America's leading QSR franchise operators, managing nearly 2,000 stores across 20 countries.

In Colombia, the company operates over 70 restaurants across 16 cities, with average monthly sales exceeding 3 million items and more than 400,000 transactions, serving approximately 50,000 customers each day.

2.1 Context: Store Operations and Management Structure

Store operations within this global QSR brand are highly standardized to ensure consistency across locations. Core elements such as equipment, suppliers, menus, pricing, and training curricula are uniformly regulated by the global franchisor. However, individual store managers retain discretion over hiring and staffing decisions. This consistency in technology and strategic practices across stores enables us to isolate variation in managerial characteristics and behavior within a stable operational environment.

2.1.1 Operations

Customers can place orders through three main channels: in-store counters, drive-throughs, and, since 2018, online delivery platforms. Once an order is placed, it is logged and automatically displayed on kitchen monitors. The kitchen is organized into five specialized stations—grill, fryer, assembly (including condiments), soda fountain, and desserts—with employees assigned to a specific station during their shift to maintain an efficient and streamlined workflow.

For example, when a burger is ordered, several team members collaborate to fulfill it. One employee toasts the bun, adds condiments and vegetables, and prepares it for assembly. Another grills the meat to standardized specifications and places it on the bun. A third completes the order by packaging the burger and adding any additional items, such as fries or drinks, before routing it to the counter, drive-through window, or delivery handoff point. In parallel, support functions—such as cleaning, equipment maintenance, security, and parking management—ensure that the overall operation runs smoothly.

2.1.2 Managerial Roles and Responsibilities

Each store is led by a store manager and supported by shift managers. Some locations also include assistant or second assistant managers. Because many core responsibilities—such as personnel oversight and inventory management—are shared across these roles, we refer to all of them collectively as "managers" throughout the study.

Managers are responsible for overseeing store operations and supervising teams of up to 80 employees working across different shifts and functional areas, including customer service, food preparation, and assembly. Their duties encompass a wide range of operational

tasks critical to store performance: calibrating kitchen equipment, ensuring sanitation standards, managing inventory and product availability, conducting quality checks, and minimizing waste. A central component of their role is personnel management, which includes scheduling, recruitment, and on-the-job training.

Managers receive extensive training to prepare them for these responsibilities and routinely rotate across stations to monitor workflow and resolve bottlenecks. Their on-the-ground presence positions them to quickly identify and address issues that could disrupt service, such as staff disengagement, equipment malfunctions, or process errors. Given their role in coordinating operations and managing staff on a daily basis, managers have a direct and substantial influence on overall store performance.

2.1.3 Consultants

Each store is also supervised by a firm consultant, who serves as a liaison between store-level management and upper management. Consultants are responsible for monitoring store performance and ensuring compliance with corporate directives, including training protocols, operational guidelines, and new technology adoption. They play a central role in our intervention: the firm designated consultants to oversee the implementation of the training curriculum and ensure that all managers in their assigned stores complete it.

2.2 Program Details and Content

We base the design of our gender-informed training curricula on the results of empirical analysis conducted in the first stage of this study, using administrative data and manager surveys from all stores in our partner firm. We estimate manager-level productivity by implementing an AKM model (Abowd et al., 1999), leveraging manager rotation across stores to isolate fixed effects representing each manager's contribution to sales per employee. We then regress these productivity estimates on a set of composite indexes derived from the manager survey. The indexes capture a range of traits and behaviors—including autonomy, rapport, decision-making, training, planning behaviors, and emphasis on metrics. To identify gender-specific patterns, we interact these indexes with manager gender and classify traits as generally important, differentially important, or uniquely predictive for male or female managers. A full description of the analysis is provided in Appendix A.

We use these findings to adapt an existing training curriculum to our context. The training intervention is based on the STITCH curriculum, originally developed to train line supervisors in the garment industry in soft skills and managerial practices linked to productivity gains (Adhvaryu et al., 2023b). The original curriculum consists of 25 in-person lessons organized into four modules[1] and has been shown to increase team productivity by 7.3 percent during implementation and 5.8 percent in subsequent months.

A condensed, online version of the curriculum—comprising 17 lessons—also exists and serves as the base for our intervention. We adapt this version to the fast-food context in Colombia by modifying examples and references to align with local norms and operational realities. We also develop nine new lessons targeting the gender-differentiated characteristics identified in our analysis that were not covered in the original content. From this expanded

[1]Figure B.1 shows the structure of the original STITCH curriculum.

pool of content, we construct two tailored curricula—one focused on tasks and metrics (male predictive) and another focused on relationships (female predictive)—selecting the lessons, both existing and new, that correspond to the traits most predictive of performance for each group. The following subsections detail the construction of the new modules and the structure of the gender-informed curricula.

2.2.1 New Lessons

We developed nine new lessons to complement the adapted STITCH curriculum. These lessons are directly informed by the empirical findings described in Appendix A and are designed to target specific behaviors and managerial characteristics found to be predictive of performance. Below, we provide a brief description of each new lesson and the behavioral index it addresses.

- *Emotions*: Builds skills in identifying emotional cues through facial expressions using exercises adapted from the "Reading the Mind in the Eyes" test (Stonewall et al., 2022). Linked to the Soft & Hard Skills index, the lesson aims to improve managers' ability to understand and respond to social signals within their teams.
- *Metrics*: Encourages managers to adopt a more strategic approach to performance metrics, focusing on selecting and interpreting key indicators relevant to their store context. Related to the Metrics index, the lesson frames metrics as decision-support tools rather than end goals.
- *Autonomy*: Helps managers assess when and how to grant autonomy to employees, as well as how to request greater decision-making space from upper management. The lesson draws on themes from both the Autonomy and Pyramid indexes.
- *Delegating*: Develops practical strategies for task delegation, enabling managers to focus on higher-leverage activities while building team capacity. This lesson addresses behaviors linked to the Pyramid and Preemptive indexes.
- *Comparative Advantage*: Introduces the concept of comparative advantage in the store context, training managers to identify role-task fit among team members. Related to the Broad Skilling index, this lesson supports more efficient task allocation.
- *Broad Skilling*: Promotes cross-training and generalist skill development among staff. Connected to the Broad Skilling index, the lesson emphasizes flexibility, resilience, and team coverage.
- *Specializing*: Presents an alternative approach to staff development, focusing on deepening expertise in specific tasks. Also tied to the Broad Skilling index, the lesson aims to help managers assess when specialization is more effective than generalization.
- *Rapport*: Offers strategies for strengthening relationships with team members, reducing interpersonal conflict, and fostering a positive store climate. This lesson is linked to the Rapport index.

- *Kaizen*: Reinforces the value of continuous improvement by teaching managers how to solicit, process, and apply feedback from staff. Based on the Kaizen index, the lesson encourages bottom-up learning and adaptive leadership.

2.2.2 New Curriculum Structure

Once the new lessons were developed, each was assigned to one of the four original STITCH modules to maintain the cohesive structure of the base curriculum. Drawing on the results from our empirical analysis (summarized in tables B.1 and B.2), both new and existing lessons were mapped to one or more behavioral indexes. Based on this mapping, lessons were classified into three groups: those included only in the task and metric curriculum, only in the relationship curriculum, or in both. Lessons that were generally predictive of performance across genders were included in both curricula, as were lessons not associated with a specific index but deemed essential for preserving the overall narrative and pedagogical flow of the program. Table 1 presents the full set of lessons, their index mappings, and their inclusion across the gender-specific curricula.

Broadly, the gender-differentiated design reflects the behavioral patterns identified in Appendix A. The task and metric curriculum emphasizes organizational structure and task allocation, with lessons focused on training strategies, delegation, and autonomy. In contrast, the relationship curriculum centers on interpersonal dynamics and workplace climate, with lessons focused on giving and receiving feedback, strengthening employee relationships, and building managerial self-confidence.

Each module includes a pre- and post-training test that all participants complete. These tests contain identical questions across both curricula and are designed to test comprehension of all module content—both shared and gender-specific. The pre-post testing framework allows us to assess attention to and retention of training content and serves as the primary outcome measure in our evaluation.

2.2.3 Design and Hosting

To deliver the intervention, each lesson was translated into Spanish and converted into a video format using the *HeyGen* video editor. Videos incorporate visuals, graphics, and examples relevant to the content, including real scenarios and quotes drawn from survey responses provided by managers at our partner firm. Each video ranges from 5 to 15 minutes in length.

The full set of lessons was hosted on the *Thinkific* platform, allowing managers to access the training as a virtual course. This format enables participants to complete the lessons at their own pace, minimizing disruption to their day-to-day responsibilities. Depending on the assigned curriculum, the total video content ranges from 2.5 to 3.5 hours.

2.3 Experimental Design

The training intervention was implemented across all 71 restaurants operated by our partner firm, involving a total of 449 managers (at the moment of randomization). Multiple managers may work within the same store and are categorized into three roles: shift managers, area managers, and store (or restaurant) managers. Randomization was conducted at the store

Table 1: New Curriculum Structure

Lesson	Mapped Index	Task/Metric Curriculum	Relationship Curriculum	Both Curricula
1. Me as a Person				
Self-esteem	Control	–	X	–
Emotions	Soft & Hard Skills	–	X	–
Behavior and values	–	–	–	X
Stress management	–	–	–	X
Showing respect	Rapport	–	–	X
Effective communication	Rapport	–	–	X
2. Me as a Manager				
My role	–	–	–	X
Plan & organize	Preemptive	–	–	X
Metrics	Metrics	–	–	X
Problem solving	Preemptive	–	–	X
Autonomy	Autonomy + Pyramid	X	–	–
Delegating	Pyramid + Preemptive	X	–	–
Conflict resolution	Rapport	–	–	X
Preventing harassment	Rapport	–	–	X
3. Me as a Team Member				
Teamwork	Pyramid + Rapport	–	–	X
Comparative advantage	Broad Skilling	–	–	X
Broad skilling	Broad Skilling	–	X	–
Specializing	Broad Skilling	X	–	–
Build accountability	–	–	–	X
Employee motivation and engagement	Pyramid + Rapport	–	–	X
4. Me as a Leader				
Rapport	Rapport	–	X	–
Give and receive feedback	Kaizen	–	X	–
Kaizen	Kaizen	–	X	–
My orientation and growth	Pyramid	X	–	–
Managing change	Personality	–	–	X
Being a leader	Rapport	–	–	X

Note: Curriculum structure. Indexes are mapped to each lesson when possible, and each lesson is assigned to either one or both of the task/metric and relationship curricula.

level to avoid spillovers and allow for assignment of a common curriculum within each store.[2]

We stratified randomization on three key variables. First, store location: 32 stores were located in Bogotá (the capital city), and 39 in other cities. Second, we used a store-level measure of manager-worker gender balance, constructed following Adhvaryu et al. (2023a). A store is considered "imbalanced" if the share of managers of a particular gender exceeds the share of frontline workers of the same gender by at least 15 percentage points. These are classified as Female-Heavy Management (FHM) or Male-Heavy Management (MHM), respectively. Balanced stores are further categorized as Male-Balanced (MB) or Female-Balanced (FB) based on which gender's representation among managers exceeds that of frontline workers by less than 15 percentage points. No stores in our sample had perfectly equal gender

[2]Ethics approval was obtained from Econometría Consultores S. A. (Consecutive Number: 011-1-2024, Date: December 27, 2024).

representation between managers and frontline workers.

We randomly assigned stores to one of three groups: a control group, a treatment group receiving the task and metric curriculum, and a treatment group receiving the relationship curriculum. Managers in the control group were given access to the hosting platform and were informed that the tests were part of a study on managerial characteristics. They did not have access to the training content and were only asked to complete the module-specific tests once, rather than at both the start and end of each module (as the treatment groups were).

To improve covariate balance across groups while preserving the integrity of randomization, we implemented a re-randomization procedure following Banerjee et al. (2020). We ran 5,000 randomization iterations and, for each, performed balance checks on baseline store-level outcomes (monthly and quarterly sales, hiring, and turnover) as well as the manager-level survey indexes described in Table A.3. We retained the top 5 percent of randomizations with the best balance profiles and selected one randomly from this subset for implementation.

Of the 71 stores, 23 were assigned to the control group, 24 to the task and metric curriculum, and 24 to the relationship curriculum. This resulted in 145 managers in the control group, 139 in the task and metric curriculum group, and 165 in the relationship curriculum group. Table B.3 presents the final assignment by treatment arm and stratification group for stores, and Table 2 for managers.

Table 2: Assignment to Treatment by Strata

	Bogotá				Other cities			
Assignment	FHM	FB	MB	MHM	FHM	FB	MB	MHM
Control	18	21	21	8	19	22	27	9
Task/Metric curriculum	10	24	20	5	16	19	30	15
Relationship curriculum	15	29	23	8	36	22	20	12

Note: Number of managers assigned to each treatment arm according to the stratification used during randomization. *FHM* is Female-Heavy Management, *FB* is Female Balanced, *MHM* is Male-Heavy Management, and *MB* is Male Balanced.

After completing the randomization process, we conducted balance tests on a set of relevant baseline variables, including both intended outcomes of the intervention and key covariates. These variables include average monthly store-level performance indicators—sales, units sold, and transaction volume—for the year prior to the intervention, as well as average monthly hiring and turnover rates over the same period. We also include the manager survey indexes used to design the training curricula.

Table 3 presents the results of the balance tests across the three experimental arms: control, task and metric curriculum, and relationship curriculum. The table reports mean values for each variable, aggregated at the store level and calculated as the average across all managers in each store.

Overall, we find no statistically significant baseline differences across groups, whether comparing each treatment to the control or the two treatment arms against each other. This

Table 3: Balance Tests Across Key Dimensions

Category	Indicator	Control vs Task/Metric	Control vs Relationship	Task/Metric vs Relationship
Performance	Avg. Monthly Sales	-7,562	2,363	12,290
		(9,181)	(4,667)	(8,119)
	Avg. Monthly Units	-4,041	-40.56	3,960
		(4,727)	(2,138)	(3,604)
	Avg. Monthly Tickets	-1,048	-131.5	784.7
		(1,372)	(627.0)	(972.9)
Personnel	Avg. Monthly Hiring	0.0882	0.131	0.174
		(0.312)	(0.135)	(0.342)
	Avg. Monthly Turnover	-0.0856	0.0373	0.160
		(0.0802)	(0.0499)	(0.106)
Survey Indexes	Control Index	-0.142	0.0226	0.187
		(0.124)	(0.0671)	(0.128)
	Personality	-0.00788	0.0647	0.137
		(0.146)	(0.0637)	(0.151)
	Experience & Education	-0.0419	0.0119	0.0658
		(0.145)	(0.0796)	(0.151)
	Soft & Hard Skills	-0.262	-0.0647	0.132
		(0.164)	(0.0747)	(0.131)
	Metrics	-0.681	-0.0389	0.603
		(0.478)	(0.132)	(0.471)
	Broad Skilling	-0.0408	0.00952	0.00026
		(0.316)	(0.154)	(0.471)
	Preemptive	-0.0156	0.00952	0.00952
		(0.154)	(0.285)	(0.220)
	Rapport	0.122	0.244	0.051
		(0.124)	(0.320)	(0.180)
	Pyramid	-0.250	0.195	-0.250
		(0.403)	(0.272)	(0.641)
	Autonomy	0.122	0.283	0.0763
		(0.407)	(0.180)	(0.403)
	Kaizen	-0.0368	-0.00575	0.0253
		(0.134)	(0.118)	(0.142)

Notes: Standard errors (clustered at the store level) in parentheses.

confirms our re-randomization procedure effectively achieved balance, supporting the internal validity of our impact estimates.

3 Data

3.1 Description

To assess knowledge acquisition across the firm, each training module utilized a standardized testing instrument administered through the Thinkific hosting platform. These tests were designed to evaluate a manager's understanding of the specific concepts taught throughout

the program and serve as the primary outcome measure for the evaluation. Crucially, every participant in the study—regardless of whether they were assigned to the control group, the relationship-focused track, or the task-focused track—responded to the exact same set of multiple-choice questions . This universal assessment included questions common to both curricula as well as questions specifically linked to the specialized content of each individual track.

The primary distinction in the testing protocol between arms was the frequency of administration. Managers in the control group were asked to complete the assessment exactly once. Because these participants did not receive any instructional content, a pre- and post-test framework was deemed conceptually unnecessary and potentially confusing for the managers. In contrast, managers assigned to either the relationship or the task and metric treatment arms completed the identical assessment twice: once as a pre-test prior to the start of the lessons and once as a post-test following the completion of the instructional videos.

Moreover, testing did not occur uniformly at the beginning or end of a fixed period: instead, participation in the platform was rolled out in an asynchronous fashion, with managers—both in treatment and control groups—completing the test (training and tests for treatment) whenever they found a suitable time within their work schedules.

A total of 238 managers completed all four modules and their corresponding tests (control group managers only completed the test). Participation in the training was voluntary, and overall completion rates at this stage of the study were approximately 60 percent.

In addition to the tests administered during the training, we conducted a baseline survey with workers who reported to participating managers. This survey was fielded prior to the start of the intervention and collected information on demographics, job satisfaction, and perceptions of managerial relationships. While this information is not used in the main analysis, balance across treatment groups (of managers) on key characteristics of their workers helps support the internal validity of the study.

The worker survey consisted of 32 questions grouped into four thematic sections: Rapport, Psychological Safety, Harassment, and Satisfaction. The Psychological Safety section included three questions that are combined into a single index, which is the measure used in our analysis. In the Harassment section, we examined each of the three questions individually. For the Rapport section, we constructed three separate indexes based on subsets of the questions. However, because not all rapport-related questions were included in those indexes—and given that the questions offered multiple response options with varying degrees of ordinality, we also conducted a principal component analysis (PCA) on all 12 rapport-related variables. This PCA incorporated a total of 35 dummy variables derived from categorical responses, from which we retained 12 principal components with eigenvalues greater than one. Finally, the Satisfaction section included questions that allowed the construction of two distinct indexes reflecting different dimensions of worker satisfaction. These composite measures serve as additional tools to assess baseline comparability and as potential mechanisms for future analysis.

Table 4: Descriptive Statistics of *Percentage Correct* by Group for Curriculum-Specific Questions

Table	Group	N	Mean	SD
Relationship-related questions	Control	115	89.88	8.66
	Relationship - pre	67	87.79	10.32
	Relationship - post	67	90.77	11.89
	Task/metric - pre	56	87.18	11.73
	Task/metric - post	56	87.82	11.40
Task/metric-related questions	Control	115	74.09	11.54
	Relationship - pre	67	72.84	12.89
	Relationship - post	67	70.75	14.91
	Task/metric - pre	56	73.57	11.51
	Task/metric - post	56	75.71	14.63
Common questions	Control	115	55.48	12.43
	Relationship - pre	67	56.62	11.21
	Relationship - post	67	68.36	15.82
	Task/metric - pre	56	57.56	11.63
	Task/metric - post	56	66.19	14.75
All questions	Control	115	66.50	9.19
	Relationship - pre	67	66.72	8.80
	Relationship - post	67	75.41	13.27
	Task/metric - pre	56	67.29	9.61
	Task/metric - post	56	73.78	12.72

Note: Descriptive statistics of percentage of correct answers for pre- and post-test questionnaires for curriculum-specific questions.

3.2 Summary Statistics

Table 4 reports descriptive statistics on the percentage of correct answers for curriculum-specific questions, disaggregated by treatment group and test timing (pre or post). For questions exclusive to each curriculum, we observe a general improvement in scores between the pre- and post-tests. Relationship-related questions show a clear increase among managers assigned to the relationship curriculum, while task- and metric-related questions show improvements among those in the task and metric curriculum. Additionally, scores on questions common to both curricula, as well as the overall set of questions, increase across both treatment groups. Questions are grouped based on the lessons to which they correspond, with those considered curriculum-specific reflecting content unique to either the relationship or task and metric track. These are listed in Appendix C.

Score variations exhibited different magnitudes based on question type. Questions exclusive to a particular curriculum showed score changes between 2.14 and 2.98 percentage points for students within that curriculum. In stark contrast, questions common to both curricula demonstrated a considerably wider variation, ranging from 8.63 to 11.74 percentage points. Across the entire set of questions, the change in scores was observed to be between 6.49 and

8.69 percentage points.

4 Treatment Effect

4.1 Training Impact on Scores

As a second step in the analysis, we compare post-test scores between treatment and control groups using the specification in Equation 1. In this model, the outcome variable Y_i denotes the share of correct answers on the post-test for manager i. The variable T_i is a dummy for the task and metric curriculum, R_i is a dummy for the relationship curriculum, and ς_s represents fixed effects for the randomization strata. Standard errors are clustered at the store level.

To account for differential compliance rates across treatment arms, we apply an inverse probability weighting procedure adapted from Wooldridge (2010). In the first step, we estimate a probit model predicting the probability of being observed (i.e., the probability of completing the post-test) on treatment indicators interacted with strata fixed effects. From this model, we generate inverse probability weights. In the second step, we estimate Equation 1 using these weights to adjust for differential completion rates.

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 R_i + \varsigma_s + \varepsilon_i \tag{1}$$

Table 5: Treatment Effects on *Percentage Correct*

	(1)	(2)
Variables	Common	All tests
Relationship curriculum	12.14***	8.290***
	(2.344)	(2.069)
Task/metric curriculum	12.21***	8.084***
	(2.576)	(2.160)
Observations	238	238
R-squared	0.178	0.137
Dep Mean Control	55.48	66.5

Note: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Effect of treatment on percentage of correct answers for post-tests. All columns include strata fixed effects, and standard errors are clustered at the store level.

Table 5 presents the average treatment effects on post-test scores across different question groupings. Columns (1) and (2) show consistent increases in the share of correct answers for both treatment groups when focusing on questions that are either common to both curricula or drawn from the full set of test items. Results are significant at the 1% level. These findings

indicate that both versions of the training effectively improved participants' understanding of the material, suggesting that the intervention successfully supported knowledge acquisition around best practices.

4.2 Training Impact on Curriculum-Specific Questions

We now turn to examining whether the training produced targeted learning effects tied to curriculum-specific content. This analysis focuses on a subset of test questions explicitly linked to topics unique to each curriculum, as detailed in Appendix C. Our identification strategy relies on the fact that while the testing instrument was identical for all managers, the instructional content they were exposed to was not. By focusing on the change in the share of correct answers (ΔY_i) between the pre- and post-test for treated managers—calculated as the difference between post- and pre-treatment scores as defined in Equation 2—we can determine if gains were concentrated in the areas emphasized by their specific curriculum. Because this outcome is only defined for individuals who received the training, the control group is excluded from this specific analysis. Since managers in one treatment arm did not receive the lessons exclusive to the other curriculum, they serve as a valid counterfactual for those specific questions while still answering the same full set of items on the post-test. The variable D_i represents either R_i (the dummy for the relationship-focused curriculum) or T_i (the dummy for the task and metric curriculum). ς_s represents randomization strata fixed effects. Standard errors are clustered at the store level, and inverse probability weights—described in prior sections—are applied.

$$\Delta Y_i = \beta_0 + \beta_1 D_i + \varsigma_s + \varepsilon_i \tag{2}$$

Table 6 reports the estimated differences in score changes between the two treatment groups, focusing on curriculum-specific questions. Column (1) examines questions specific to the relationship curriculum and uses R_i as the independent variable, while column (2) examines questions specific to the task and metric curriculum and uses T_i as the independent variable. The remaining columns, which examine questions common across both curricula (column (3)) and all test questions (column (4)) use R_i as the independent variable.

For these outcomes, the group assigned to the alternate curriculum serves as a valid counterfactual. This is because managers from one treatment arm do not receive the lessons exclusive to the other curriculum, yet they respond to the same full set of post-test questions. Therefore, any differential improvement in performance on curriculum-specific questions can be attributed to exposure to the corresponding training content. This setup strengthens the internal validity of the comparison between treatment groups.[3]

The results of this analysis strengthen the internal validity of the study by showing that learning outcomes aligned precisely with the content of the assigned track. Managers in the task and metric group demonstrated a 4.8 percentage point improvement on questions spe-

[3]The only exception is Question 6 from Module 3, which covers material that appears in both curricula. However, the lesson framing differs between the two, leading to a different "correct" answer for those taking the relationship curriculum and those taking the task and metric curriculum. All other targeted questions are linked to lessons that are exclusive to a single curriculum.

Table 6: Treatment Effects on Post - Pre Difference

Variables	(1) Relationship	(2) Task/metric	(3) Common	(4) All tests
Treatment Difference	4.879		1.988	1.722
(Relationship - Task/Metric)	(3.017)		(2.499)	(1.764)
Treatment Difference		4.786*		
(Task/Metric - Relationship)		(2.691)		
Observations	123	123	123	123
R-squared	0.060	0.077	0.012	0.011
Dep Mean	1.92	-.16	10.33	7.69

Note: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Differential treatment effect for curriculum-specific questions. All columns have strata fixed effects and standard errors are clustered at the store level. One singleton observation is dropped.

cific to their curriculum compared to those in the relationship group, while relationship-track managers improved by 4.9 percentage points on their specific content relative to their counterparts. Because the test itself never changed, these differential improvements can be attributed solely to the exposure to corresponding training material. Furthermore, the lack of significant differences in scores for common questions and overall test scores suggests that specialized training does not compromise the acquisition of general managerial best practices. Taken together, these findings underscore the value of aligning training content with the concepts most relevant to different manager profiles without sacrificing broader knowledge acquisition.

Taken together, these results demonstrate that specialized training curricula can lead to measurable learning gains in targeted skill areas. While both curricula effectively support broader knowledge acquisition, the presence of curriculum-specific learning differences highlights the potential for tailored instructional designs to build specific managerial competencies. These findings suggest that when training is designed to address specific skill sets—whether focused on interpersonal dynamics or operational metrics—it can effectively reinforce those underdeveloped areas while still maintaining a baseline of general managerial knowledge.

5 Conclusion

This study provides novel evidence on the benefits of tailoring managerial training to reflect diverse approaches to leadership. While women remain underrepresented in managerial positions in the service sector, prior research suggests that stores led by female managers often outperform those led by male managers—particularly through stronger team cohesion, improved scheduling, and more responsive leadership.

Our findings highlight that tailoring training content to specific managerial skills can lead to measurable learning gains in those targeted areas. Managers assigned to each curriculum showed greater improvements in the content aligned with their assigned track, while

no meaningful differences were observed in more general content. This suggests that when training is designed to address skill areas that may be underdeveloped or underemphasized, it can help close those gaps. For example, managers exposed to task- and metric-focused content—areas historically less emphasized in the firm's training—displayed significant gains in related knowledge. These targeted gains do not appear to come at the expense of more general learning: both curricula also led to improvements in common content, indicating that differentiated training can enhance specific competencies while still reinforcing core managerial knowledge. These patterns highlight the potential of tailored training strategies to strengthen diverse skill sets and support managers with varied starting points.

These findings support a move away from one-size-fits-all training strategies. Recognizing that effective management can take different forms—and that these differences correspond to measurable practices and outcomes—offers a pathway for firms to optimize training and promotion pipelines. By investing in programs that reflect the diversity of effective management styles, organizations can promote more inclusive leadership development while also enhancing overall performance and productivity. At the same time, we acknowledge the importance of avoiding essentialist interpretations that might inadvertently reinforce gender-based stereotypes. The training curricula were not targeted by gender, and the current analysis does not examine heterogeneous effects by gender identity. Future work, enabled by newly available data on the gender of store managers, will examine whether the impacts of training vary systematically across gender lines. This will allow us to assess whether the tailoring of training content aligns with or challenges prevailing patterns in managerial effectiveness across genders and to do so in a way that remains attentive to equity and gender stereotypes.

Looking ahead, the follow-up administrative data collection will be crucial in evaluating the extent to which gains in knowledge and curriculum-specific learning lead to sustained improvements in managerial effectiveness and store performance. By linking test results to operational and workforce outcomes, we aim to identify whether targeted training translates into better-managed teams and more efficient store operations. Moreover, by revisiting workers' survey responses, we will assess whether the training generates perceptible changes in how employees experience and evaluate their managers. This next stage of the analysis will provide a deeper understanding of the practical implications and long-run benefits of investing in tailored management training.

References

Abowd, J. M., Creecy, R. H., and Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data. Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J., and Zilibotti, F. (2007). Technology, information, and the decentralization of the firm. *The Quarterly Journal of Economics*, 122(4):1759–1799.

Adhvaryu, A., Howell, P., Nyshadham, A., and Tamayo, J. A. (2023a). Rapport in organizations: Evidence from fast food. Working Paper 24-032, Harvard Business School.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2018). The skills to pay the bills: Returns to on-the-job soft skills training. Working Paper 24313, National Bureau of Economic Research.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2022). Management and shocks to worker productivity. *Journal of Political Economy*, 130(1):1–47.

Adhvaryu, A., Murathanoglu, E., and Nyshadham, A. (2023b). On the allocation and impacts of managerial training. Working Paper 31335, National Bureau of Economic Research.

Adhvaryu, A., Nyshadham, A., and Tamayo, J. A. (2023c). Managerial quality and productivity dynamics. *The Review of Economic Studies*, 90(4):1569–1607.

Aghion, P., Bloom, N., Lucking, B., Sadun, R., and Van Reenen, J. (2021). Turbulence, firm decentralization, and growth in bad times. *American Economic Journal: Applied Economics*, 13(1):133–169.

Banerjee, A. V., Chassang, S., Montero, S., and Snowberg, E. (2020). A theory of experimenters: Robustness, randomization, and balance. *American Economic Review*, 110(4):1206–1230.

Battistelli, A., Montani, F., Guicciardi, M., and Bertinato, L. (2016). Regulation of exercise behaviour and motives for physical activities: The Italian validation of BREQ and MPAM-R questionnaires. *Psychologie Française*, 61(4):333–348.

Beam, E. A., Merfeld, J. D., and Wickremeratne, N. (2025). Improving gender norms in the workplace. Pre-results paper.

Block, J. and Kremen, A. M. (1996). IQ and ego-resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology*, 70(2):349–361.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., and Van Reenen, J. (2019). What drives differences in management practices? *American Economic Review*, 109(5):1648–1683.

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

Bloom, N., Lemos, R., Sadun, R., Scur, D., and Van Reenen, J. (2021). World management survey - Manufacturing. Harvard Dataverse.

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

Bloom, N. and Van Reenen, J. (2011). Human resource management and productivity. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4B, pages 1697–1767. Elsevier.

Boissiere, M., Knight, J. B., and Sabot, R. H. (1985). Earnings, schooling, ability, and cognitive skills. *The American Economic Review*, 75(5):1016–1030.

Borghans, L., Duckworth, A. L., Heckman, J. J., and ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Bowles, S., Gintis, H., and Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, 39(4):1137–1176.

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, 117(1):339–376.

Ellison, S. F. and Snyder, C. M. (2014). An empirical study of pricing strategies in an online market with high-frequency price information. Working Paper 4655, CESifo.

Feldberg, A. C. (2022). The task bind: Explaining gender differences in managerial tasks and performance. *Administrative Science Quarterly*, 67(4):1049–1092.

Frederiksen, A., Kahn, L. B., and Lange, F. (2020). Supervisors and performance management systems. *Journal of Political Economy*, 128(6):2123–2187.

Friebel, G., Heinz, M., Hoffman, M., Kretschmer, T., and Zubanov, N. (2025). Is this really kneaded? Identifying and eliminating potentially harmful monitoring practices. Working Paper 34122, National Bureau of Economic Research.

Gibbons, R. and Waldman, M. (2004). Task-specific human capital. *American Economic Review*, 94(2):203–207.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1):26–42.

Goldin, C. (2015). Hours flexibility and the gender gap in pay. Report, Center for American Progress.

Goldin, C. and Katz, L. (2011). The cost of workplace flexibility for high-powered professionals. *The Annals of the American Academy of Political and Social Science*, 638(1):45–67.

Groves, T., Hong, Y., McMillan, J., and Naughton, B. (1994). Autonomy and incentives in Chinese state enterprises. *The Quarterly Journal of Economics*, 109(1):183–209.

Haegele, I. (2021). The broken rung: Gender and the leadership gap. Working Paper. arXiv:2404.07750.

Halim, S. E. and Mansyur (2023). How motivation affects productivity in the workplace. *Jurnal Manajemen Bisnis*, 10(1):304–318.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

Hoffman, M. and Tadelis, S. (2021). People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of Political Economy*, 129(1):243–285.

Jovanovic, B. (1979). Firm-specific capital and turnover. *Journal of Political Economy*, 87(6):1246–1260.

Judge, T. A., Piccolo, R. F., and Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, 89(1):36–51.

Korman, A. K. (1966). "Consideration," "initiating structure," and organizational criteria: A review. *Personnel Psychology*, 19(4):349–361.

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In Lefcourt, H. M., editor, *Research with the Locus of Control Construct*, volume 1: Assessment Methods, pages 15–63. Academic Press, New York.

Martin, A. J. (2005). The role of positive psychology in enhancing satisfaction, motivation, and productivity in the workplace. *Journal of Organizational Behavior Management*, 24(1-2):113–133.

Mas, A. and Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, 107(12):3722–3759.

McKenzie, D. and Woodruff, C. (2016). Business practices in small firms in developing countries. *Management Science*, 63(9):2967–2981.

Metcalfe, R. D., Sollaci, A. B., and Syverson, C. (2023). Managers and productivity in retail. Working Paper 31192, National Bureau of Economic Research.

Mincer, J. A. (1974). Introduction to "schooling, experience, and earnings". In *Schooling, Experience, and Earnings*, pages 1–4. National Bureau of Economic Research.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics*, 13(4):653–677.

Peterson, C. and Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. Oxford University Press.

Ranganathan, A. and Shivaram, R. (2021). Getting their hands dirty: How female managers motivate female worker productivity through subordinate scut work. *Management Science*, 67(5):3299–3320.

Reinhardt, F., Alvarez, J. B., Buitrago, J. M., and Cal, M. (2020). Arcos Dorados: A QSR recovery plan. Case 721-023, Harvard Business School. Revised April 2021.

Reis, R. (2006). Inattentive producers. *The Review of Economic Studies*, 73(3):793–821.

Rosenberg, M. (2015). *Society and the Adolescent Self-Image*. Princeton University Press.

Sinclair, V. G. and Wallston, K. A. (2004). The development and psychometric evaluation of the Brief Resilient Coping Scale. *Assessment*, 11(1):94–101.

Stogdill, R. M. and Coons, A. E., editors (1957). *Leader Behavior: Its Description and Measurement*. Bureau of Business Research, Ohio State University.

Stonewall, J. H., Ouverson, K. M., Helgerson, A., Gilbert, S. B., and Dorneich, M. C. (2022). An evaluation to determine if Reading the Mind in the Eyes scores can be improved through training. *PLOS One*, 17(4):e0267579.

US DOC (2015). Management and Organizational Practices Survey (MOPS). U.S. Census Bureau.

Vlachopoulos, S. P. and Michailidou, S. (2006). Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: The Basic Psychological Needs in Exercise Scale. *Measurement in Physical Education and Exercise Science*, 10(3):179–201.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2nd edition.

A Stylized Facts

A.1 Data: Manager Survey

Our analysis relies heavily on a novel survey that we conducted which encompasses the entire population of store and shift managers employed by our research partner from August 2022 to February 2023. Table A.1 provides an overview of our sample. During our study period, we surveyed 438 managers, of whom 48 percent were female. Sixty-four percent of our sample are shift managers, 21 percent are assistant or second assistant store managers, and 15 percent are restaurant managers.

Table A.1: Survey Sample

	Count	% Total
Number of Respondents	438	100%
Female Managers	210	48%
Manager Types		
Restaurant Manager	66	15%
Assistant Manager	92	21%
Shift Manager	280	64%

The survey comprised the following topics:

- **Attention:** This area reflects the contributions from five managerial practices: efforts to achieve production targets, the frequency of production monitoring, active personnel engagement, communication, and overcoming worker resistance. The first element evaluates the number of practices a manager engages in to ensure production deadlines are met. The second records the number of times per day that a manager makes rounds to identify production-related problems. The third item measures the number of practices a manager engages with to retain workers, motivate low-performing ones, and support high-performing subordinates. The fourth element captures the methods managers use to convey production-related issues with both workers and upper-level managers. The last element addresses how often managers face reluctance from subordinates to change or innovate.

  Within the literature, these features are called managerial attention, with some recent evidence coming from Reis (2006); Ellison and Snyder (2014) as well as from Adhvaryu et al. (2022), who found that more attentive managers are more capable of identifying and relieving production bottlenecks that might arise from shock to worker productivity.

- **Autonomy:** We evaluate this item via two approaches. The first is the one suggested by Stogdill and Coons (1957) (i.e., the Ohio State Leadership model) who generated two managerial styles: Consideration and Initiating structure. The former captures the degree to which managers play an active role in guiding tasks within teams, whereas the latter captures the degree of good rapport with subordinates (Korman, 1966).

According to Judge et al. (2004), this classification has also been consistently validated as a successful measure of leadership.

The second approach relies on the methodology suggested by Vlachopoulos and Michailidou (2006), which gathers information about the degree of alignment between workers´ preferences and the scope of decision-making and autonomous behavior within job roles. With this, we can capture managers' autonomy in terms of both leadership style and management practices.

Autonomy has a widespread effect throughout the organizational hierarchy (Groves et al., 1994), and given the notion of middle-range that managers have within our implementing partner, it is a crucial aspect of managerial quality that should be fostered (Bloom and Van Reenen, 2011). In turn, this could lead to more empowered managers and more decentralized processes, which are key elements to enhancing resilience during economic deceleration (Aghion et al., 2021). According to Bresnahan et al. (2002), productivity returns might be high when management is decentralized, and Acemoglu et al. (2007) suggests this might be achievable when companies are closer to the technological frontier.

- **Cognitive skills:** Productivity and earnings appear to be highly correlated with cognitive skills (Boissiere et al., 1985; Bowles et al., 2001). To inform this element, we included two math-related tests. The first is a 16-question assignment to assess the number of correct answers and the performance rate. The second is a 9-sequence progressive test that evaluates short-term memory skills.

- **Control:** Risk aversion, patience, and locus of control are also measured in the instrument. The latter module is designed to capture the degree to which a participant believes in their ability to control events or elicit changes under specific circumstances (Levenson, 1981).

- **Demographics:** To determine the sample composition, the survey classifies participants by gender, age, job title, and level of education.

- **Motivation:** Productivity seems closely related to the degree of worker engagement and their motivation to complete tasks (Halim and Mansyur, 2023; Martin, 2005). To capture this component, we present a 15-item module that breaks down total motivation into four subcategories, following the approach of Battistelli et al. (2016). The first category describes the extent to which individuals rely heavily on the effect of external elements to drive their decision-making processes (i.e., external motivation), and the second category (introjected regulation) describes the level by which the motivation of the behavior has been internalized. High levels describe a negative internalized motivation by which individuals do activities. The third element—identified regulation—captures the behavior of an individual where their actions are driven by reasons personally important to them, whereas the fourth element—intrinsic regulation—captures the fact of doing a particular activity for the inherent satisfaction that can be extracted from it, rather than for some subsequent consequence.

- **Personality:** Psychometric and personality traits have begun to receive particular attention in the literature, along with the holistic notion of the workers (i.e., cognitive skills and human capital accumulation) given their impact on productivity (Borghans et al., 2008; Heckman and Kautz, 2012). Specifically, we included some modules for conscientiousness, extroversion, agreeableness, emotional stability, and openness following the methodology of Goldberg (1992). Additional sections also address self-esteem (Rosenberg, 2015), industriousness (Peterson and Seligman, 2004), and resilience (Sinclair and Wallston, 2004; Block and Kremen, 1996).

- **Tenure:** Human capital accumulation is key for productivity at both the industry and firm levels (Mincer, 1974; Jovanovic, 1979; Neal, 1995; Gibbons and Waldman, 2004), Hence, we assess job experience by measuring the number of years working in the QSR industry, the types of previous positions held in the sector, and whether participants have worked in other industries before.

- **World management:** The information provided by Bloom et al. (2021) in the World Management Survey was adapted to the local environment in which QSR firms operate. This allows us to identify the frequency of common practices such as product prioritization and sales forecasting and to understand the perception of challenges to meet both long-term and short-term targets.

- **Management practices and context-specific:** Approaches to define and monitor key performance indicators (KPI) are measured under the guidance of US DOC (2015), adapted to the local context, as well as to understand the managerial preferences to have specialized workers and hire new ones when demand shocks affect store performance. Additional elements, such as continuous improvement and strategies to deal with simultaneous responsibilities, are also measured.

Table A.2 outlines the modules included in the survey, providing the names of each module, the number of questions per section, and the standard deviations and means. The mean represents the average success rate on the tests (correct answers as a percentage of total questions). For example, the average performance rate in the math test was 91%, whereas that for the digit recall test was 43%. The Job Description and Job Satisfaction sections are on a 1–7 Likert scale, with 1 being the lowest satisfactory level and 7 the highest one. The remaining modules are presented in a 1–5 Likert agree scale. For instance, the mean of 3.83 in the personality traits module suggests that the average participant, within the sample, has a balanced approach to discovering new things and reaching goals.

The variables in the survey are recorded as ordered categorical, select one, select multiple, and open-ended data input questions. The ordered categorical variables and open-ended numerical questions are standardized to facilitate the comparison of coefficients across regressions, ensuring a mean of 0 and a standard deviation of 1 for the survey measure. Select-one and select-multiple questions have been transformed into binary variables, where 1 indicates that the manager selected a particular option and 0 indicates otherwise.

Following these data transformations, we have 350 variables capturing various manager characteristics and behaviors. To better categorize characteristics and behaviors that can predict store performance, we created indexes from some survey modules and then grouped

Table A.2: Summary Statistics by Module

Module	Mean	SD	No. of questions
Personality Traits	3.83	.58	7
Conscientiousness	4.50	.53	10
Extraversion	3.75	.71	10
Agreeableness	4.42	.52	10
Emotional Stability	3.72	.59	10
Openness	4.19	.51	10
Locus of Control	3.73	.77	5
Self-Esteem	4.44	.56	10
Industriousness	4.52	.52	8
Autonomy	4.46	.60	4
Resilience	4.55	.54	4
Motivation	4.14	.47	15
Job Description	5.78	.78	6
Job Satisfaction	6.15	.88	12
Leadership Style (Consideration)	3.24	.39	10
Leadership Style (Initiating structure)	3.65	.43	10
Ravens Test (Success rate)	.52	.21	9
Reading the Eyes Test (Success rate)	.65	.12	36
Math Test (Success rate)	.91	.12	16
Digit Recall Test (Success rate)	.43	.27	9

Notes: The variables across the survey modules are recorded in various formats, including ordered categorical variables answered on a 1–5 Likert scale (except for Job Description and Job Satisfaction which are on a 1–7 Likert scale), select-one and select-multiple questions, and open-ended data input questions. To facilitate comparison of coefficients across regressions, the ordered categorical variables and open-ended numerical questions are standardized, ensuring a mean of 0 and a standard deviation of 1 for the survey measure.

indexes and questions according to the characteristics they best capture. Afterward, we ran factor analysis on these groups and created new indexes that describe manager behaviors according to their survey answers. Table A.3 shows the names and a brief description of each of the 11 indexes that were created. These indexes allow us to have fewer variables that capture specific characteristics and behaviors in order to identify better those that significantly affect manager and store productivity. In order to have more comparable measures, all indexes are standardized.

A.2 Manager Productivity

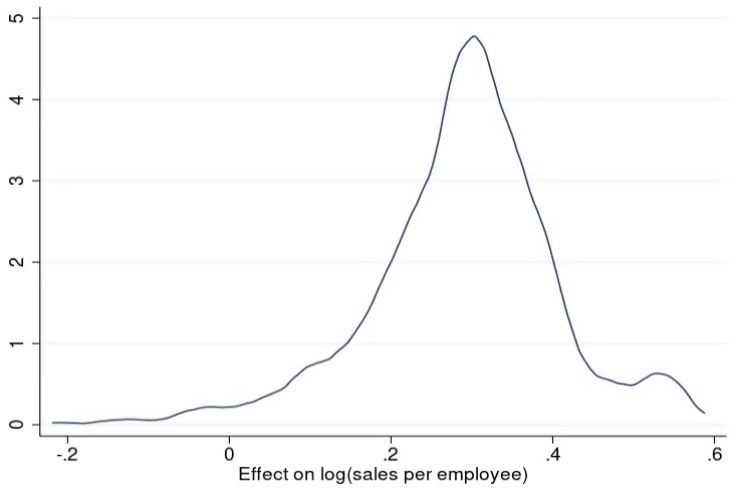
In our initial analysis, we examine which manager survey measures are most predictive of manager performance and how this relationship varies based on the manager's gender. To do this, we first calculate the contribution of each manager to store performance. By leveraging the rotation of managers between stores for reasons unrelated to performance, we can estimate the differences in store productivity that can be attributed to individual managers. To do this, we implement an AKM model following Abowd et al. (1999) with the outcome variable being the logarithm of monthly sales per employee.

We estimate the following two-way fixed effects model:

$$Y_{ijt} = \theta_i + \Psi_{J(i,t)} + \delta_t + \varepsilon_{ijt} \tag{3}$$

where the dependent variable Y_{ijt} is the monthly sales per employee in store j under manager i in month t. $\Psi_{J(I,t)}$ refers to a store fixed effect, and δ_t is a date (monthly) fixed effect. As discussed in Abowd et al. (1999, 2002), manager fixed effects are identified only within *connected sets* of stores, linked by the movements of managers across stores. We identify four such connected sets and estimate equation (3) within each connected set.

Figure A.1: Distribution of Manager Fixed Effects in Log(Sales per employee)



Notes: The figure shows the manager fixed effects distribution from estimating equation (3). We use Log(Sales per employee) as the outcome variable.

Figure A.1 shows the distribution of the manager fixed effects. Once we have obtained the fixed effect of each manager, we estimate the effect of each of the previously created indexes

Table A.3: Final Indexes Description

Index Name	Description
Control	Measures control, resilience, and self-esteem. A higher value implies the manager has scored higher in those modules.
Personality	Measures personality traits. A higher value implies more conscientiousness, agreeableness, openness, industriousness, and autonomy.
Experience & Education	Measures experience and education. A higher value implies more years of experience in the fast-food sector and more education.
Soft and Hard Skills	Measures skills. A higher value implies better scores in the Ravens, Reading the Mind in the Eyes, digit-span recall, and arithmetic tests.
Metrics	Measures attention to metrics. A higher value implies KPIs and other metrics are kept and checked more often and more rigorously.
Broad Skilling	Measures generalization in employee training. A higher value implies employees are trained in more stations and keep a broader skill set, and a lower score implies employees are more specialized in particular skills and stations.
Preemptive Planning	Measures planning capacity. A higher value implies the manager does planning tasks directly, such as sales forecasting, turn assignment, supply and expiry date checks, and product prioritization, among others.
Rapport	Measures managers' relationship with employees. A higher value implies fewer conflicts with employees, more attention to star performers, and more attention to employees' motivation and satisfaction.
Pyramid Organization	Measures compliance with a pyramidal organization structure. A higher value implies more delegating of tasks to employees, as well as reliance on employees for problem-solving, and less reliance on upper management for responsibilities and problem-solving.
Autonomy	Measures autonomy. A higher value implies more autonomy from managers, including taking steps independently for problem-solving and less involvement of upper management on store activities.
Kaizen	Measures the taking and implementing of feedback. A higher value implies managers receive more feedback from employees and implement feedback from employees more often.

Notes: The indexes were created using factor analysis grouping questions related to similar topics and behaviors.

based on the survey on manager productivity. To do that, we first do a simple OLS approach regressing manager performance on survey indexes, and afterward, to identify differential effects, we employ a simple cross-sectional difference-in-differences style approach, regressing manager performance on gender, survey indexes, and their interaction. We estimate the following model:

$$Y_i = \beta_0 + \beta_1 Characteristic_i^k + \mu_i \tag{4}$$

$$Y_i = \beta_0 + \beta_1 female_i + \beta_2 Characteristic_i^k + \beta_3 Characteristic_i^k * female_i + \mu_i \tag{5}$$

Y_i is the measure of productivity, namely the fixed effect of manager i. $female_i$ is a binary variable that takes a value of 1 when the manager is female and 0 otherwise. $Characteristic_i^k$ is manager i's survey index k.

We are especially interested in the second specification, as it allows us to identify the effects of each index on managers of each gender. To do that, we consider three estimates: β_2, β_3, and $\beta_2 + \beta_3$, from which we are able to look at characteristics that matter for store productivity when managers are male (β_2) and female ($\beta_2 + \beta_3$) and the differential impact between men and women (β_3).

For this initial analysis, we categorize the significant variables into four classes: those generally predictive of performance, those differentially predictive of performance, those predictive of performance for male managers only, and those predictive of performance for female managers only.

Variables are sorted into groups based on the value and significance of relevant estimates:

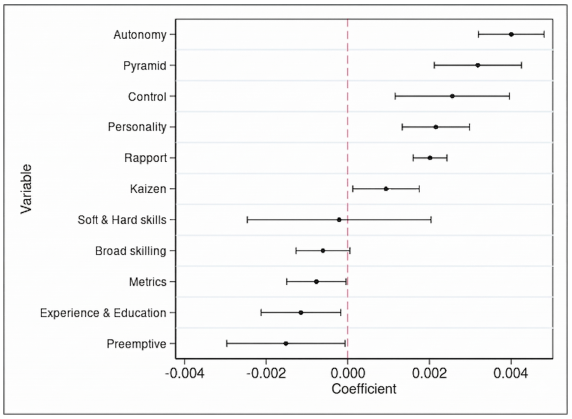
- *Generally important:* β_2 and $\beta_2 + \beta_3$ are significant and have the same sign
- *Differentially important:* β_2 and $\beta_2 + \beta_3$ are significant and have opposite signs
- *Male important only:* β_2 is significant and $\beta_2 + \beta_3$ is insignificant
- *Female important only:* $\beta_2 + \beta_3$ is significant and β_2 is insignificant

Figures A.2a and A.2b display the estimated values with 95% confidence intervals for each index. Following the figures, there seems to be no *generally important* index, as no estimate is significant with the same sign for both men and women. There are two *differently important* indexes: Soft & Hard skills and Broad skilling. The estimates for these indexes imply that higher scores in skill tests are positively correlated with female productivity and negatively correlated with male productivity. In a consistent pattern, training employees to be generalists—captured by the Broad Skilling index—is also positively correlated with productivity for women and negatively correlated for men. These results suggest that while certain practices like cross-training and high cognitive or "soft" skill scores are associated with higher performance for female managers, they may be negatively associated with the performance of their male counterparts within this specific institutional context.

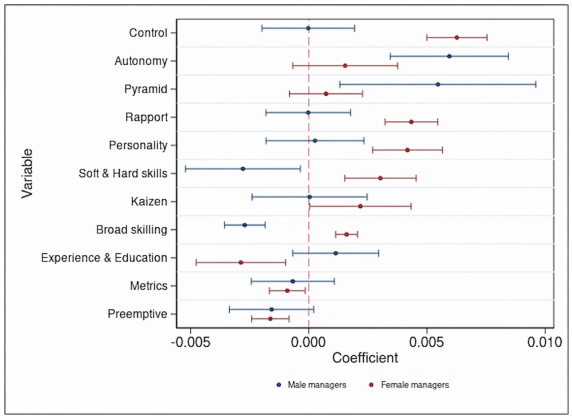
Two indexes can be considered *male important*, Autonomy and Pyramid, which imply that male managers' productivity is correlated with being more independent from upper management for decision-making and problem-solving and with relying on employees for both

Figure A.2: Effect of survey indexes individually on manager productivity

(a) Effect on all managers



(b) Effect by manager gender



Notes: Figure A.2a shows the estimates from equation (4). Figure A.2b shows the estimates from equation (5). All standard errors are clustered at the connected set level.

decision-making and problem-solving. On the other side, seven indexes can be considered *female important*, four positively: Control, Rapport, Personality, and Kaizen, and three negatively: Experience, Metrics and Preemptive. This implies that female managers' productivity is positively correlated with having higher self-esteem and control, having positive personality traits such as agreeableness and openness, having better relations with employees, and listening and implementing feedback from employees, and is negatively correlated with more planning and keeping on metrics.

This initial exercise, however, is limited to each index exclusively and does not take into account how each index can interact with the others nor how a fuller profile of manager behavior can interact with manager productivity by taking into account all indexes at the same time. To take these problems into consideration, we estimate a second set of models by aggregating all indexes in the same equation:

$$Y_i = \beta_0 + \sum_{k=1} \beta_k Characteristic_i^k + \mu_i \tag{6}$$

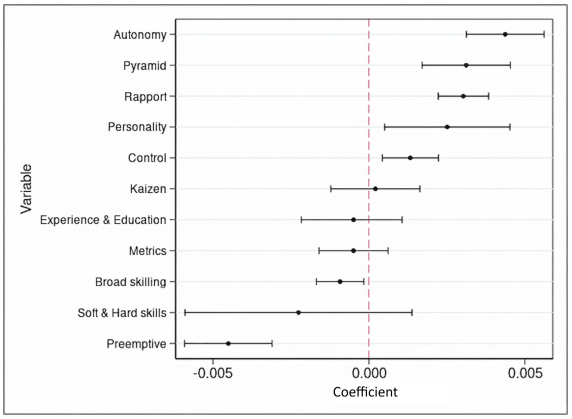
$$Y_i = \beta_0 + \beta_1 female_i + \sum_{k=1} [\gamma_k Characteristic_i^k + \omega_k Characteristic_i^k * female_i] + \mu_i \tag{7}$$

In this case, variables are sorted according to their relevant estimates again:

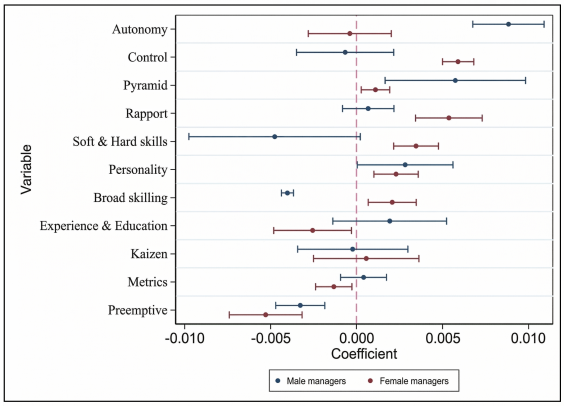
- *Generally important:* γ_k and $\gamma_k + \omega_k$ are significant and have the same sign
- *Differentially important:* γ_k and $\gamma_k + \omega_k$ are significant and have opposite signs
- *Male important only:* γ_k is significant and $\gamma_k + \omega_k$ is insignificant
- *Female important only:* $\gamma_k + \omega_k$ is significant and γ_k is insignificant

Figure A.3: Effect of all survey indexes on manager productivity

(a) Effect on all managers



(b) Effect by manager gender



Notes: Figure A.3a shows the estimates from equation (6). Figure A.3b shows the estimates from equation (7). All standard errors are clustered at the connected set level.

Figures A.3a and A.3b display the estimated values with 95% confidence intervals for each index. Following the figures, there appear to be three *generally important* indexes: Pyramid Organization and Personality, which are both positively correlated with performance for both men and women, and Preemptiveness, which is negatively correlated for both groups. There is one *differently important* index: Broad skilling. While training employees to be generalists—as captured by the Broad Skilling index—is positively correlated with productivity for female managers, it is negatively correlated for male managers. These results confirm that even when controlling for a full profile of managerial behaviors, the performance implications of cross-training strategies remain significantly differentiated by the manager's gender.

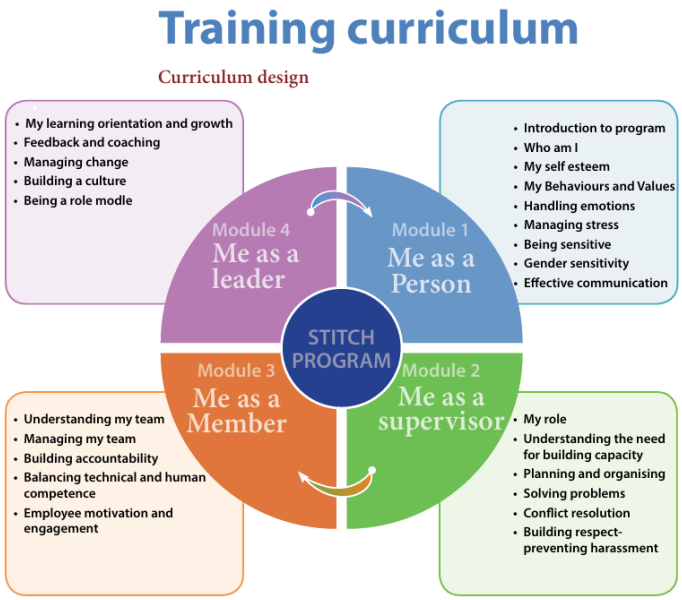
One index can be considered *male important*, Autonomy, which implies that male managers' productivity correlates with being more independent from upper management for decision-making and problem-solving. Conversely, five indexes can be considered *female important*: three positively (Control, Rapport, and Soft & Hard skills) and two negatively (Metrics and Experience). This implies that female managers' productivity is positively correlated with having higher self-esteem and control, having better relations with employees, and achieving higher scores on skill tests. Conversely, their productivity is negatively correlated with putting more relevance on metrics and checking them more often. These results suggest that while Soft & Hard skills are positively associated with female productivity, they are not a significant predictor of performance for male managers in this setting.

When considering equations (5) and (7) jointly, there appear to be two *generally important* indexes: Personality, which is positively correlated with performance for both men and women, and Preemptiveness, which is negatively correlated for both groups. There is one *differently important* index: Broad skilling. While training employees to be generalists—as captured by the Broad Skilling index—is positively correlated with productivity for female managers, it is negatively correlated for male managers. These results confirm that even when controlling for a full profile of managerial behaviors, the performance implications of cross-training strategies remain significantly differentiated by the manager's gender.

Two indexes can be considered *male important*, Autonomy and Pyramid Organization. This implies that male managers' productivity correlates with being more independent from upper management for decision-making and relying more on a pyramidal delegation structure within their teams. Conversely, five indexes can be considered *female important*: three positively (Control, Rapport, and Soft & Hard skills) and two negatively (Metrics and Experience). This implies that female managers' productivity is positively correlated with having higher self-esteem and control, having better relations with employees, and achieving higher scores on skill tests. Conversely, their productivity is negatively correlated with a heavy reliance on metrics and higher levels of previous industry experience.

B Additional Figures and Tables

Figure B.1: Original STITCH modules and lessons



Notes: Structure of the original STITCH training curriculum.

Table B.1: Sign of the Effect of Indexes on Male Managers

Index	Productivity	
	(Single indexes)	(All indexes)
Control		
Personality		+
Experience & Education		
Soft & Hard Skills	−	
Metrics		
Broad Skilling	−	−
Preemptive Planning		−
Rapport		
Pyramid Organization	+	+
Autonomy	+	+
Kaizen		

Notes: Maps the sign of all significant effects of indexes on male managers' productivity.

Table B.2: Sign of the Effect of Indexes on Female Managers

Index	Productivity	
	(Single indexes)	(All indexes)
Control	+	+
Personality	+	+
Experience & Education	−	−
Soft & Hard Skills	+	+
Metrics	−	−
Broad Skilling	+	+
Preemptive Planning	−	−
Rapport	+	+
Pyramid Organization		+
Autonomy		
Kaizen	+	

Notes: Maps the sign of all significant effects of indexes on female managers' productivity.

Table B.3: Assignment to Treatment by Strata

	Bogotá				Other cities			
Assignment	FHM	FB	MB	MHM	FHM	FB	MB	MHM
Control	3	4	3	1	3	3	4	2
Task/Metric curriculum	3	3	3	1	4	3	5	2
Relationship curriculum	3	4	3	1	4	3	4	2

Note: Number of stores assigned to each treatment arm according to the stratification used during randomization. *FHM* is Female-Heavy Management, *FB* is Female Balanced, *MHM* is Male-Heavy Management and *MB* is Male Balanced.

Table B.4: Balance Test, Workers Survey. Control vs Relationship Curriculum

Variable name	Mean Control	SE Control	Mean Relationship	SE Relationship	Diff.	p-value
safety_index	6.48	0.12	6.32	0.12	-0.16	0.4454
harassment_1	0.07	0.02	0.07	0.02	0.00	0.9748
harassment_2	0.01	0.01	0.02	0.01	0.01	0.2136
harassment_3	0.02	0.01	0.03	0.01	0.01	0.5842
rap_collab_index	12.52	0.25	12.04	0.25	-0.49	0.2804
rap_clinician_input	5.74	0.14	5.27	0.14	-0.47	0.0768 *
rap_Reverse_Clinician	6.09	0.13	5.76	0.13	-0.33	0.1499
rap_full_index	24.35	0.41	23.09	0.41	-1.27	0.1239
PC1	-0.22	0.13	0.20	0.14	0.42	0.1642
PC2	0.19	0.11	-0.14	0.11	-0.34	0.0903 *
PC3	0.01	0.09	0.05	0.09	0.05	0.7874
PC4	-0.08	0.09	-0.01	0.09	0.07	0.7134
PC5	0.05	0.09	-0.01	0.09	-0.05	0.6847
PC6	-0.09	0.08	0.14	0.08	0.23	0.0818 *
PC7	-0.08	0.08	0.10	0.09	0.18	0.1831
PC8	0.03	0.08	0.03	0.08	0.00	0.9715
PC9	0.04	0.08	0.03	0.07	-0.01	0.9045
PC10	0.04	0.06	-0.02	0.08	-0.07	0.5368
PC11	-0.01	0.07	0.05	0.07	0.07	0.4914
PC12	-0.03	0.07	0.02	0.07	0.05	0.6628
PC13	0.00	0.07	0.03	0.07	0.03	0.7946
PC14	0.00	0.07	0.02	0.07	0.02	0.8487
satis_agreement	9.81	0.16	9.58	0.15	-0.23	0.4476
satis_feeling	3.22	0.07	3.13	0.07	-0.09	0.4754

Balance test for the workers survey comparing those stores in the control group to those stores in the relationship curriculum treatment.

Table B.5: Balance Test, Workers Survey. Control vs Task/Metric Curriculum

Variable name	Mean Control	SE Control	Mean Task/Metric	SE Task/Metric	Diff.	p-value
safety_index	6.48	0.12	6.29	0.12	-0.19	0.3405
harassment_1	0.07	0.02	0.04	0.01	-0.02	0.4989
harassment_2	0.01	0.01	0.03	0.01	0.02	0.0803 *
harassment_3	0.02	0.01	0.04	0.01	0.02	0.2239
rap_collab_index	12.52	0.25	12.11	0.24	-0.41	0.4058
rap_clinician_input	5.74	0.14	5.45	0.13	-0.29	0.231
rap_Reverse_Clinician	6.09	0.13	5.86	0.12	-0.23	0.2576
rap_full_index	24.35	0.41	23.46	0.38	-0.89	0.2828
PC1	-0.22	0.13	0.02	0.13	0.24	0.3531
PC2	0.19	0.11	-0.05	0.10	-0.24	0.224
PC3	0.01	0.09	-0.06	0.10	-0.07	0.7387
PC4	-0.08	0.09	0.09	0.09	0.17	0.4227
PC5	0.05	0.09	-0.04	0.09	-0.09	0.4953
PC6	-0.09	0.08	-0.05	0.09	0.03	0.8041
PC7	-0.08	0.08	-0.02	0.08	0.06	0.6702
PC8	0.03	0.08	-0.06	0.08	-0.09	0.4686
PC9	0.04	0.08	-0.07	0.08	-0.11	0.3338
PC10	0.04	0.06	-0.02	0.08	-0.07	0.5602
PC11	-0.01	0.07	-0.04	0.08	-0.02	0.8414
PC12	-0.03	0.07	0.01	0.08	0.04	0.7021
PC13	0.00	0.07	-0.03	0.07	-0.03	0.8072
PC14	0.00	0.07	-0.02	0.07	-0.01	0.9095
satis_agreement	9.81	0.16	9.60	0.16	-0.21	0.4841
satis_feeling	3.22	0.07	3.24	0.06	0.02	0.9109

Notes: Balance test for the workers survey comparing those stores in the control group to those stores in the task/metric curriculum treatment.

Table B.6: Balance Test, Workers Survey. Relationship Curriculum vs Task/Metric Curriculum

Variable name	Mean Relationship	SE Relationship	Mean Task/Metric	SE Task/Metric	Diff.	p-value
safety_index	6.32	0.12	6.29	0.12	-0.04	0.8576
harassment_1	0.07	0.02	0.04	0.01	-0.02	0.4524
harassment_2	0.02	0.01	0.03	0.01	0.01	0.5807
harassment_3	0.03	0.01	0.04	0.01	0.01	0.5851
rap_collab_index	12.04	0.25	12.11	0.24	0.07	0.8739
rap_clinician_input	5.27	0.14	5.45	0.13	0.18	0.4657
rap_Reverse_Clinician	5.76	0.13	5.86	0.12	0.10	0.6482
rap_full_index	23.09	0.41	23.46	0.38	0.37	0.6465
PC1	0.20	0.14	0.02	0.13	-0.18	0.542
PC2	-0.14	0.11	-0.05	0.10	0.09	0.5927
PC3	0.05	0.09	-0.06	0.10	-0.11	0.5793
PC4	-0.01	0.09	0.09	0.09	0.10	0.6111
PC5	-0.01	0.09	-0.04	0.09	-0.03	0.784
PC6	0.14	0.08	-0.05	0.09	-0.20	0.1653
PC7	0.10	0.09	-0.02	0.08	-0.12	0.361
PC8	0.03	0.08	-0.06	0.08	-0.09	0.4869
PC9	0.03	0.07	-0.07	0.08	-0.10	0.4305
PC10	-0.02	0.08	-0.02	0.08	0.00	0.9988
PC11	0.05	0.07	-0.04	0.08	-0.09	0.3606
PC12	0.02	0.07	0.01	0.08	-0.01	0.951
PC13	0.03	0.07	-0.03	0.07	-0.07	0.5931
PC14	0.02	0.07	-0.02	0.07	-0.04	0.7442
satis_agreement	9.58	0.15	9.60	0.16	0.02	0.9518
satis_feeling	3.13	0.07	3.24	0.06	0.11	0.4375

Balance test for the workers survey comparing those stores in the relationship curriculum treatment to those stores in the task/metric curriculum treatment.

C Curriculum-Focused Questions

C.1 Relationship-focused questions

Module 1 – Me as a Person

1. Having high levels of self-esteem will help you stay calm in unexpected situations.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

2. Emotions are not consciously controlled, and thus our body can reflect when we feel strong emotions.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

3. Our emotions have an impact on our behaviour and energy levels.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

4. Emotions are tied to immediate reactions, rather than thought out responses.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

Module 3 – Me as a Team Member

1. To make the most out of your stores, employees should be trained:
   a) In many different stations, so they have a broad understanding
   b) In just a few stations, so they can specialize
   c) Only in the stations they enjoy working at
   d) I don't know

2. Employees that can work in several stations can empathize better with their colleagues.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

Module 4 – Me as a Leader

1. Having a good relationship with your employees can help with communication and productivity in your store.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

2. Which of the following is important when building good relationships?
   a) Affinity
   b) Empathy
   c) Rudeness
   d) I don't know

3. Finding common ground and shared experiences with your employees can help foster good relationships.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

4. Giving negative feedback spoils relationships between people and hampers effective working.
   ☐ Yes (1) ✓ **No (2)** ☐ I don't know (9)

5. Listening to feedback from your employees makes finding and solving problems easier.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

C.2 Task/metric-focused questions

Module 2 – Me as a Manager

1. A Manager should only delegate tasks that involve low effort and can be easily accomplished.
   ☐ Yes (1) ✓ **No (2)** ☐ I don't know (9)

2. Knowing which tasks employees can do by themselves without supervision and can be delegated to them is important for managing a restaurant.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

Module 3 – Me as a Team Member

1. To make the most out of your stores, employees should be trained:
   a) In many different stations, so they have a broad understanding
   b) In just a few stations, so they can specialize
   c) Only in the stations they enjoy working at
   d) I don't know

2. Employees who are specialized in one station can serve as mentors to other employees.
   ✓ **Yes (1)** ☐ No (2) ☐ I don't know (9)

Module 4 – Me as a Leader

1. Which of the following helps me to be open to learning and growing continuously .
   a) Fixed mind-set
   b) Growth mind-set
   c) Social mind-set
   d) Business mind-set
   e) I don't know