I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17727

Managing Skills in Organizations - Evidence from a Field Experiment

Leonhard Grabe
Dirk Sliwka

FEBRUARY 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17727

Managing Skills in Organizations - Evidence from a Field Experiment

Leonhard Grabe
University of Cologne

Dirk Sliwka
University of Cologne, CESifo and IZA

FEBRUARY 2025

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Managing Skills in Organizations - Evidence from a Field Experiment*

We study the value of skill management in organizations. In a natural field experiment with 2,582 service technicians, we vary managers' ability to monitor and manage employee skills. We find that removing managers' access to hard information on employee skills reduced training intensity, work performance, and job satisfaction. Combining detailed personnel records and survey data, we show that the intervention lowered employee efforts to identify training needs and managerial attention to employee development. In particular, high-skill employees received less training to broaden their skill set and, in turn, performance losses are driven by higher completion times for complex work assignments.

JEL Classification: J24, J28, M12, M53

Keywords: skill management, training, performance, field experiment

Corresponding author:
Dirk Sliwka
University of Cologne
Faculty of Management, Economics and Social Sciences
Albertus-Magnus-Platz
50923 Köln
Germany

E-mail: sliwka@wiso.uni-koeln.de

* We thank Jakob Alfitian, Robert Dur, Florian Engelmaier, Matthias Heinz, Christian Hofmann, Saskia Opitz, Christopher Stanton, Matthias Sutter, as well as participants of the 2023 Field Days Conference, the 2023 Advances in Field Experiments Conference, the 2023 Annual Meeting of the Society for Industrial and Organizational Economics and the 2024 CESIfo Venice Summer Institute for helpful comments. L.G. received an internship contract from the firm to facilitate access to confidential data. The firm has exerted no influence on the data analysis or the scientific conclusions drawn in this paper. This study has been pre-registered on OSF (ID: AUWM2). IRB approval was obtained by the German Society for Experimental Economic Research (GfeW). Funding by the German Research Foundation (DFG) under Germany's Excellence Strategy – EXC 2126/1 – 390838866 is gratefully acknowledged.

1 Introduction

Human capital is a critical driver of firm performance, which is why firms have a vested interest in a highly-skilled workforce (Bartel, 1989; Black & Lynch, 1996; Bloom, Sadun, & Van Reenen, 2010). To benefit from ongoing technological innovations, firms must ensure that their employees are prepared for the ever-evolving work requirements (Acemoglu & Restrepo, 2022; Autor, Chin, Salomons, & Seegmiller, 2024; Brynjolfsson, Mitchell, & Rock, 2018). As acquiring new skills is central to this transformation in many industries, understanding and developing employee skills become crucial for firm success. Historically, measuring and monitoring employee skills has been challenging and costly. However, the widespread adoption of learning technologies has significantly lowered the barriers to systematic skill management.[1] While there exists a large literature in labor economics on the strong link between employee skills, performance and earnings (Acemoglu & Autor, 2011; Alfonsi, Bassi, Rasul, & Spadini, 2024; Bapna, Langer, Mehra, Gopal, & Gupta, 2013; Hanushek, Schwerdt, Wiederhold, & Woessmann, 2015; Mincer, 1974), and a growing literature on management practices that are conducive to employee development (Buell, Cai, & Sandino, 2022; Cai, Chen, & Chen, 2024; Friebel, Heinz, & Zubanov, 2022; Sandvik, Saouma, Seegert, & Stanton, 2024; Sandvik, Saouma, Seegert, & Stanton, 2020), the internal mechanisms by which firms invest into employee skills remain under-explored.

We make use of a unique setting which allows us to study a skill management process in detail using a rich set of information on skill assessments, training participation, supervisor interactions as well as productivity data and survey results. In a natural field experiment involving 2,582 field service technicians from a large technology company, we vary managers' ability to monitor and manage employee skills. The skill management process of the firm is as

[1]Despite the importance of skills in the labor economics, the literature shows little consensus on a definition of employee skills. The OECD (2017) differentiates between cognitive, non-cognitive, and technical skills. In this paper, we focus on technical skills, which are "combinations of cognitive and non-cognitive skills used to accomplish specific tasks" (Brunello & Wruuck, 2021, p. 1146). Therefore, skill management includes all activities to facilitate an employee's development and use of skills in the interest of the firm.

follows: Employees in the organization are regularly invited take part in online assessments on job-specific technical skills. Managers are instructed to hold quarterly development meetings where they review the results of these assessments and also discuss training recommendations made by the employee. Based on the assessment results and the discussion with the employee, the manager then decides which training the employee receives and when. In our experiment, managers (and their respective teams) are randomly allocated to a control or treatment group. Prior to the intervention and in the control group, employees were encouraged to share the outcomes of their skill assessments with their managers.[2] In the treatment group, employees were told not to report assessment outcomes and managers were instructed not to demand them. Employees in both groups were encouraged to recommend training measures in the meetings and managers had to decide about training assignments. In other words, by removing managers' access to hard information on employees' skills, we reduce their ability to actively manage these skills. This change allows us to examine how active skill management affects the allocation of training and the work performance of employees.

Our key results are the following: Reducing managers' ability to monitor and manage employee skills had crucial consequences for employee development, as it led to a substantial reduction in the overall likelihood of receiving training. Moreover, the intervention also reduced employee performance, increasing the time needed to complete key work tasks. A back-of-the-envelope calculation suggests that the negative effect on work performance would – when workforce size were to be adapted accordingly – be associated with about 1,040,000 USD higher personnel costs per annum. Moreover, contrary to a large strand of research claiming detrimental effects of monitoring on psychological well-being, we find that the treatment also significantly reduced employee job satisfaction. Hence, a reduction in monitoring not only had no positive effect on employee satisfaction – on the contrary: we find that regular

[2]Assessment outcomes include the date, subject and total score of each test, as well as the question text and answer selected by the employee. Therefore managers primarily learn two things from the report: a) whether or not the employee has invested effort in to discovering his or her own training needs and b) the current skill level as measured by the assessment itself.

monitoring of skills assessments is associated with higher employee well-being.

To investigate the underlying mechanisms in more detail, we study the role of employees and managers in the skill management process. On the employee side, we show that the treatment lowered employee efforts to identify their own training needs, as it reduced their propensity to take part in the skill assessments when participation is no longer monitored. Evidence from the post-experimental survey further shows that some employees are particularly motivated to achieve a good result in the assessment to receive appreciation from their managers and that this motivation is reduced when assessment scores are no longer revealed to managers. Our results thus underscore the key role of supervisor monitoring in organisations as a means to increase reputational incentives.

On the side of the managers, we show that the intervention reduced attention towards employee development as it reduced the likelihood that supervisors arrange development meetings with their employees. A detailed analysis of the treatment effects on training and team performance reveals that this shift in manager behavior has strong implications for the firm's overall performance. While the intervention moderately reduced training assigned to address skill gaps among low-skilled employees, it in particular affected the more highly skilled employees where the reduction in meeting frequencies and, in turn, also of assigned trainings is more pronounced. Although a reduction of human capital investments for more highly skilled employees may appear less problematic at first glance, a deeper analysis reveals that the reduction is particularly significant in trainings aimed at broadening these employees' skill sets. Consequently, we find that the overall losses in work performance induced by the treatment were primarily driven by longer completion times for complex work assignments. These results thus highlight the importance of skill management not only to overcome skill gaps among low-skill employees but also in promoting continuous training and upskilling among high-skilled employees.

Our paper contributes to the study of training within firms. While there is a growing literature that quantifies the direct and indirect returns of training in organizations (Adhvaryu, Nyshadham, & Tamayo, 2023; De Grip & Sauermann, 2012; Espinosa & Stanton, 2023), "little work has covered how firms decide on training investments" (Hoffman & Stanton, 2024, p. 78). Employees often take up training on their own initiative, which can be inefficient, as those who are most likely to gain from training are the least likely to take it up (Delfino, Garnero, Inferrera, & Sadun, 2024; Sandvik et al., 2024). Therefore, the standard rationale behind the managerial assignment of training is to focus on low- or under-performing employees with the goal of raising performance up to that of the average employee (Adhvaryu, Kala, & Nyshadham, 2022; Diaz, Neyra-Nazarrett, Ramirez, Sadun, & Tamayo, 2024). In contrast to this approach, we find that investing more time and effort into broadening the skill set of high-performing employees can be more fruitful in an environment that is characterized by ongoing technological change.

Our paper further contributes to the research on management practices (Bloom, Sadun, & Van Reenen, 2007, 2012; Gosnell, List, & Metcalfe, 2020). Several studies have shown that monitoring employees is positively related to firm-level productivity and profitability[3], as it solves classical moral hazard problems (Holmström, 1979; Merchant & Van der Stede, 2017). On the other hand, the introduction of a monitoring technology is often perceived critically by employees, potentially leading to lower work morale, feelings of distrust and higher levels of stress.[4] We show, however, that the monitoring of skills is not only beneficial for productivity but also for employees' job satisfaction, not only by helping to overcome employees' reluctance to discover skill gaps, but also by fostering managerial support in the process of addressing them. Finally, our results also contribute to the literature on career and image

[3]For instance, Jackson & Schneider (2015) show that moral hazard explains why auto mechanics are reluctant to use check lists to monitor their performance, giving managers in a retail bank access to objective performance measures increases profits (Manthei & Sliwka, 2019), randomly installing GPS tracking devices in a Nigerian trucking firm increases average speed (Rochambeau, 2019), increasing the visibility of the monitoring technology improves task performance in some dimensions (Jensen, Lyons, Chebelyon, Bras, & Gomes, 2020), and introducing performance review meetings in a retail chain raises profits (Manthei, Sliwka, & Vogelsang, 2023).

[4]See for instance Falk & Kosfeld (2006), Christ (2013) , Nebeker & Tatum (1993), and Ravid, White, Tomczak, Miles, & Behrend (2023)

concerns in organizations (Bénabou & Tirole, 2006; Holmström, 1982) by providing evidence on employees' motivation to signal their capabilities to their superiors. Prior research shows that in situations of (managerial) observation, people manipulate signals of skill (Burks, Carpenter, Goette, & Rustichini, 2013; Ewers & Zimmermann, 2015) and concentrate their effort on tasks that allow them to showcase their ability (Campbell, Epstein, & Martinez-Jerez, 2011; De Janvry, He, Sadoulet, Wang, & Zhang, 2023; Katok & Siemsen, 2011). While these phenomena have primarily been documented in laboratory settings, field (experimental) evidence on signaling in the workplace is scarce.

The paper proceeds as follows: Section 2 illustrates the operating environment of the firm and their skill management process. Section 3 lays out the experimental manipulation and shows descriptive statistics of key dependent variables. In Section 4 we show causal effects of the intervention on trainings, work performance, and employee well-being. We then analyze potential mechanisms to understand the role of managers and employees in the skill management process. In Section 6 we show how these mechanisms affect training and work performance. Section 7 concludes.

2 The Setting

We run the experiment in the technical service division of a large service organization. The organization employs service technicians operating in 249 teams, 125 of which are part of this experiment. Service technicians mostly work independently and their primary responsibility is to install and maintain the firm's products and services. While some assignments can be performed remotely, most of the work is done by one technician at the client site, i.e. in private homes or business offices. Within each team, every employee is responsible for a designated geographical area. Employees receive their daily work schedules from a regional dispatcher. The team manager coordinates the content of the work schedule with the dispatcher and ensures that the technicians on their team have the skills required to complete their assignments within the allotted time frame (see Figure A1 for an overview of the relationship between employee, manager and regional dispatcher). Given the technical nature of the work tasks and ongoing technological changes in the industry, continuous skill management is essential to maintaining high levels of productivity and customer satisfaction. An overview of the firm's skill management process is presented in Figure 1. The first step in this process involves employees participating in regular knowledge assessments to evaluate their skills. Assessments are conducted bi-annually for each of six core technical skill domains (hereinafter referred to as "skills"). Each employee has a skill profile specifying a subset of of the six core skills, and these skill profiles determine the work assignments that can be allocated to them. Most employees have two or three skills and they are automatically invited to the respective assessments. Hence, a typical employee is supposed to take part in 4-6 skill assessments per year. Supervisors may also assign new skills to employees in order to broaden their skill sets. Skill assessments are administered through multiple-choice questions on an online platform, accessible via tablet, phone, or laptop. Employees receive an automated invitation and are given a designated one-hour time frame during which no work assignments are scheduled. Each assessment contains 20 questions, randomly selected from a pool of 60 questions per skill. If an employee answers a question incorrectly, the correct answer, along with a brief explanation, is provided before they proceed to the next question. All questions are created by experienced technicians and reviewed by the works council to ensure that they cover the actual work content appropriately.

After completing the assessment, employees receive an overall rating of their current skill level, along with a detailed record of their responses to every question. They can then choose from a set of suggested training measures, which they may recommend directly to their manager. During this step, employees can also write a personal message that will be transmitted along with their training proposals (the corresponding screen in the learning tool is shown in Section A.2). In the following step, managers and employees discuss the results of the skill

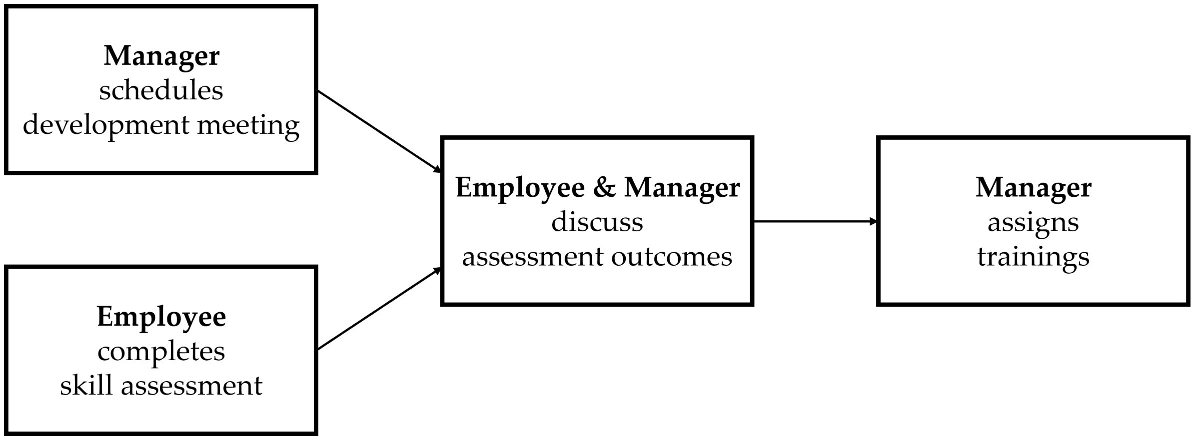

Figure 1: Overview of the skill management process of the firm.

assessment and the employee's training proposals in the development meeting. Trainings are classified into those targeted at filling gaps in skills employees are already supposed to possess ("skill preservation") and trainings that broaden employees' skills ("new skill acquisition"), for instance learning to work on tasks based on new technologies that will become more important in the future or are needed to solve more complex and non-routine problems. It is thus possible that two employees receive training on the same broader topic, but the content differs because the training has a different purpose. Given the technical nature of the work, most trainings take place in person. In these training sessions, employees practice work on real products with their daily tools under the supervision of an instructor.

Before the intervention and in the control group, employees are asked to bring their evaluation sheets to the development meeting. Managers then decide on the allocation and scheduling of training sessions, based on the employee's proposals and the results of the skill assessment. Since the introduction of the learning platform, the firm's management and its works council have been debating the appropriate level of monitoring within the organization.[5] The

[5]In Germany, employees have a right to set up an employee-elected works council in establishments with more than 5 employees. Firms need the consent of the works council when implementing policies to monitor employees' performance. § 87 (1) of the Works Constitution Act [BetrVG], for instance specifies that "The works council has a right of co-determination [..] in the introduction and use of technical devices designed to monitor the behavior or performance of the employees".

works council argues that showing assessment detailed results to managers is unnecessary and potentially degrading, suggesting that employees might be discouraged from participating in assessments due to fear of being blamed for poor performance. To gain a deeper understanding of these concerns, we conducted ten short phone interviews with five supervisors and five employees from different teams, who were referred to us by the firm. In these interviews, some employees indeed expressed that fear of being blamed for poor outcomes might contribute to a reluctance to participate in skill assessments or to share their results in the development meeting. Although the firm's leadership has historically been in favor of monitoring employee outcomes, they were open to reducing it if doing so would improve assessment participation and encourage training, particularly among low-skill employees.

3 Experimental Setup

We randomly assigned 125 teams (N = 2,582 employees) to either a treatment or the control group. In both groups, employees were encouraged to propose potential training measures to their managers. In the treatment group, employees were explicitly instructed *not* to share their assessment outcomes with their managers, and managers were told not to request them, but rather to focus on the training measures proposed by their employees. Treatment assignment was stratified at the regional level. We further included team size, average team assessment participation and assessment scores as stratification variables.[6] The experiment was conducted from July 1st 2022, until March 31st 2023. Employees in both groups were informed about upcoming changes to the learning platform via email and directly through the platform when completing a skill assessment (see Figure A2). All managers and employees were invited to participate in virtual Q&A sessions, which were split by treatment group and role. The post-experimental survey was distributed via email on behalf of the CEO in January 2023. All participants who completed the survey received a compensation of €7 via bank

[6]The firm is organized in six regional areas that comprise a total of 37 regions. Stratified randomization was performed in Stata using the package *randtreat* (https://acarril.github.io/posts/randtreat).

Table 1: Balance Table

	Control (N=63)		Treatment (N=62)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Assessment Participation	0.81	0.16	0.80	0.17	-0.02	0.03
Assessment Score (Std.)	0.00	0.28	0.00	0.34	0.00	0.06
Number of Trainings	5.03	6.03	4.85	6.21	-0.18	1.10
Supervisor Meetings	0.86	0.34	0.78	0.41	-0.07	0.08
Troubleshooting (in min.)	34.41	3.79	34.05	4.14	-0.36	0.73
Customer Service (in min.)	7.19	1.27	6.68	0.98	-0.51	0.21
Quality Control (in min.)	10.34	2.08	10.54	2.28	0.20	0.40
Documentation (in min.)	4.87	0.36	5.02	0.48	0.15	0.08
Team Size	20.00	2.16	20.13	2.80	0.13	0.45

Note: Table shows pre-treatment data from all 125 teams included in the experiment. Assessment participation denotes the average participation for the assessments every member of the team has been assigned to. Assessment results indicate the average standardized ratio of correctly answered questions. Trainings denote the average number of trainings a supervisor has assigned to the team. Supervisor meetings is a team average of the number of development meetings between an employee. Troubleshooting, Customer Service, Quality Control and Documentation denote the weekly average completion time for a given task across all members of a team.

transfer.[7] Unless stated otherwise, we use all available data from the beginning of 2022 until the end of the post-experimental survey in March 2023. For the data analysis, we excluded employees who changed teams before the end of March 2023, as they may have been part of different treatment groups. Table 1 provides summary statistics for relevant pre-experimental outcomes for the treatment and control group. In addition to the stratification variables – assessment participation, standardized assessment score, and team size – we also consider key dependent variables for the analysis, such as the number of training measures and the average duration of key work tasks.[8] Prior to the intervention employees on average took part in about 80% of the skill assessments relevant to them and attended about 0.5 trainings per year.

[7]This particular survey was conducted by the researchers and respondends were informed that the company would only receive aggregated responses and anonymity is preserved. Compensation for survey participation was paid from university research funding.

[8]Note that it was not possible to measure the number of meetings on an individual level prior to the experiment. Instead, data collection for this time period was limited to the team level.

4 Results

Our main results are presented in three parts. First, we document main treatment effects of the intervention on employee training on the aggregate. Second, we investigate the downstream effect on team performance, as measured by the weekly average time employees need to complete the most common work tasks. Finally, we quantify the effects on employee well-being.

4.1 Trainings

First, we examine the effect of the treatment on employee development, as measured by the number of trainings assigned by the supervisor. The average treatment effect on the number of trainings is depicted in Figure 2 and the corresponding regression results are displayed in Table 2, where we also analyze the likelihood of an employee receiving at least one training in the treatment period. In columns (2) and (4), we control for the pre-treatment participation rate and the pre-treatment assessment scores.[9] As Figure 2 shows, the number of trainings per employee in the first 6 months of the treatment period is reduced by 0.19 compared to the control group average of 0.37, which corresponds to a 51.89% reduction (p = 0.036). Estimating the extensive margin effect – i.e. the likelihood of receiving any training at all – we find that employees who are not allowed to bring their assessment results to the development meeting are 7.23 percentage points less likely to receive any training in the first 6 months of the treatment period, which translates into a 46.85% reduction (p = 0.03) in comparison to the control group (see columns 3 and 4 of Table 2). Therefore, reducing managers' ability to directly monitor and manage employee skills substantially reduces training intensity and the overall likelihood of receiving training.

[9]In Table A1, we use replicate our results using a Pseudo-Poisson ML estimator for the treatment effects in columns (1) and (2) of Table 2 and a logit model to estimate the treatment effects depicted in columns (3) and (4). We further re-estimate the treatment effect on the number of trainings on the team-level where we can control for pre-treatment values and target a longer time period (12 months instead of 6), which also replicates our results.

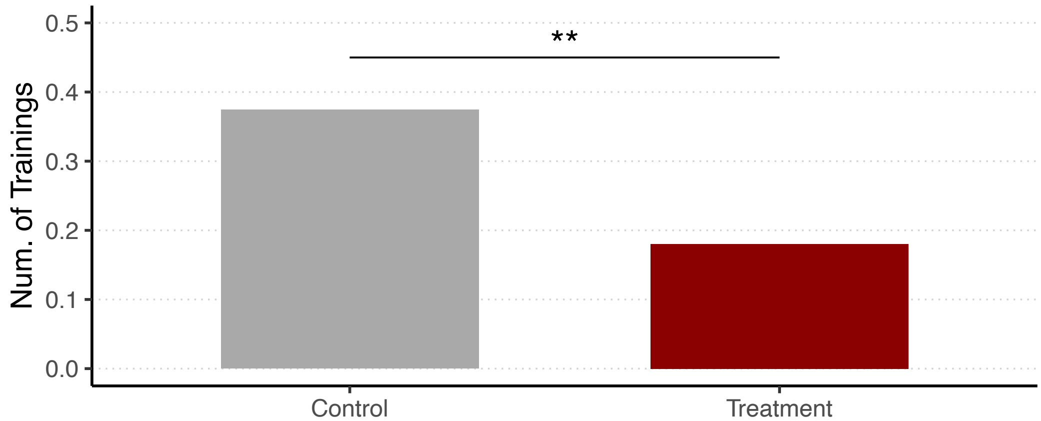

Figure 2: Treatment Effect on the Number of Trainings per Employee.

Table 2 reveals two additional observations. First, prior assessment participation is correlated with subsequent training activity. That is, while those employees who never participated in a skill assessment prior to the experiment have about 0.13 trainings on average; this number increases to 0.32 for those technicians who participated in every assessment they were assigned priorly (see column 2). This supports the assumption that assessment participation generally matters for training assignment. Second, we find that prior assessment scores are weakly positively correlated with the likelihood of receiving training, but not the overall number of trainings. We explore the underlying mechanisms that drive the reduction in training in Section 5.

Table 2: Treatment Effects on Trainings

	Num. of Trainings		Prob. of Training	
	(1)	(2)	(3)	(4)
Treatment	−0.194**	−0.219**	−0.072**	−0.077**
	(0.091)	(0.103)	(0.033)	(0.036)
Prior participation		0.321**		0.139***
		(0.129)		(0.043)
Prior score		0.021		0.012
		(0.019)		(0.008)
Constant	0.374***	0.131	0.154***	0.046
	(0.077)	(0.112)	(0.027)	(0.041)
Observations	2582	2078	2582	2078

Note: Table shows estimates from linear regressions on the number of assigned trainings per employee and the likelihood of receiving at least one training assignment. Columns (1) and (2) use the number of trainings as a dependent variable, which is the sum of all trainings associated with key skills. In columns (3) and (4) the dependent variable is a binary indicator for receiving training which takes value 1 if an employee has received at least one training and 0 otherwise. Prior participation captures the participation rate prior to the experiment. Prior score denotes the average standardized ratio of correctly answered questions on a skill assessment. Robust standard errors clustered on the team level in parentheses. **, *** indicate significance on the 5% and 1% level, respectively.

4.2 Work Performance

To understand the downstream effects of the intervention on employee performance, we utilize the firm's performance tracking system, which measures the time a technician takes to complete a task in a given work assignment. We focus on four work phases that are common for almost every work assignment, accounting for 95% of all observations in the data set.[10] Customer services (M = 6.27 min) includes all tasks associated with customer interaction, primarily prior to the appointment. For most clients, employees are required to call the customer to understand their infrastructure and their current issue. This short conversation often serves as a precursor to the on-site troubleshooting. Troubleshooting (M = 43.29 min) is the core of a technician's job and involves all tasks associated with problem-solving at the client site. Troubleshooting is the most time-intensive task where technical knowledge and training are most likely to play a significant role. Quality control (M = 10.22 min) follows troubleshooting and involves assessing the client's infrastructure, typically after resolving the issue. If any further problems remain, this serves as documentation for a follow-up visit. Finally, documentation involves all tasks associated with paperwork or data entry (M = 4.48 min). We have access to detailed team-level data on task execution times for all customer visits. Our data set comprises teams weekly average time spent on these four task phases, measured in minutes. When calculating the weekly average time per task, the system automatically trims the top and bottom 5% of work assignment durations each week. The data set covers every week from half a year before until one full year after the treatment was introduced. For the analysis, we perform difference-in-differences estimations for the average time needed to complete each step in the work assignment. We include team fixed effects and assignment-specific time trends to account for seasonal changes in product sales (e.g. holiday seasons).[11] As Table 3 shows,

[10] One row in the data set represents the average time to complete a specific activity in a given week across all members of one team. The tasks presented in Table 3 cover 95% of all rows in the data set. The remaining tasks concern for instance the delivery of equipment to a work site or manual programming.

[11] A detailed analysis regarding the nature of the work assignment follows in Section 6, where we further differentiate work assignments by their degree of complexity.

Table 3: Treatment effects on average time per task

	Customer Service	Troubleshooting	Quality Control	Documentation
	(1)	(2)	(3)	(4)
Treatment	0.509**	1.381*	−0.051	−0.033
	(0.252)	(0.766)	(0.462)	(0.100)
Team FE	Yes	Yes	Yes	Yes
Type × Week FE	Yes	Yes	Yes	Yes
Mean Dep. Var	6.27	43.29	10.22	4.48
Observations	33 880	42 843	34 734	38 191

Note: Table shows average treatment effects from difference-in-difference regressions on the average time per task in minutes. All models include team as well as Type × Week fixed effects. Assignment types are usually associated with a specific technology, but do vary with respect to their degree of complexity. Averages from the control group are depicted below the estimates. Standard errors clustered on the team-level in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

reducing the access to employee skill assessments reduces employee performance, as evident by the increased average assignment times in two key phases. Column (1) shows that teams in the treatment group need 0.51 minutes longer to complete the same customer service assignments as teams in the control group. In comparison to the control group mean prior to experiment, the treatment thus caused an 8.11% increase in the average time spent on customer service tasks ($p = 0.046$). The average time spent on troubleshooting activities increases by 1.38 minutes (see column 2) or 3.19% ($p = 0.074$). We do not find effects on the last two steps of the work assignment, such as quality control (column 3) and documentation (column 4). A back-of-the-envelope calculation illustrates that the personnel costs associated with the treatment can be substantial. Assuming that employees complete two customer visits per day that include troubleshooting and customer service, the estimated time increase of the treatment is equal to 33500 hours per year. This corresponds to the annual working hours of 20 full-time employees or about 1 million USD in estimated personnel costs.[12]

[12] In this calculation, we neither consider the time saved by the reduction in training, nor potential relative decrease in wages that stems from the reduction in skill acquisition. Therefore, our calculations should be considered as an upper bound. Wage information was retrieved from the German union for technical workers.

4.3 Employee Well-Being

While most research in economics documents positive effects of monitoring on performance, a large strand of psychological research suggests that increased levels of monitoring can reduce employee well-being (Deci, Connell, & Ryan, 1989; Nebeker & Tatum, 1993; Ravid et al., 2023) – a point also often emphasized in the popular press (Kantor & Sundaram, 2022; Murty & Karanth, 2022; Shrikant, 2023). These findings could imply that removing managers' ability to directly observe the outcomes of the skill assessment reduced the perceived pressure and in turn enhanced job satisfaction in our context as well. To study the effect on employee well-being, we use data from the post-experimental survey, linking it with skill assessment and training data (see Section A.14 for all survey items). The overall participation in the post-experimental survey is 36.76%. While survey participation was larger in the control group, participants in both groups exhibit very similar characteristics with respect to prior participation in skill assessments and achieved assessment scores (for further details see Table A3). Analyzing treatment effects on psychological outcomes shows a universally negative impact: As shown in Table 4, the treatment significantly reduced job satisfaction (column 1, survey item "*All things considered, I am quite satisfied with my current job.*") and supervisor support (column 2), which refers to the perceived attention and effort supervisors invest in their employees' development (seven items, e.g. "*My supervisor is interested in what I'm currently learning.*"). As column (3) shows, the treatment does not appear to reduce employees' turnover intention. It is important to recall that in the control group employees had the same access to the skill assessments and also were encouraged to state their training needs. Nevertheless, the active management of skills through closer scrutiny promoted employee well-being, a finding we explore in more detail below.

Table 4: Treatment effects on employee perceptions

	Job Satisfaction	Supervisor Support	Turnover Intention
	(1)	(2)	(3)
Treatment	−0.200**	−0.214**	0.088
	(0.077)	(0.097)	(0.079)
Constant	0.094*	0.098	−0.040
	(0.053)	(0.062)	(0.055)
Observations	775	848	819

Note: Table shows estimates from linear regressions on standardized measures of job satisfaction, supervisor support and turnover intentions as reported by the employees in the post-experimental survey. Job satisfaction and turnover intentions are measured using single-item measures from the Linked Personnel Panel (Kampkoetter et al., 2016), our measure of supervisor support is based on a four-item survey from Decius, Schaper, and Seifert (2021) and a three-item scale from Udris and Riemann (1980). All questions are answered on a 7-point likert scale from 1 = 'strongly disagree' to 7 = 'strongly agree'. Robust standard errors clustered on team in parentheses, *, ** indicate significance on the 10% and 5% level, respectively.

5 Mechanisms

Having established that reducing managers' ability to actively manage employee skills negatively impacts training intensity, performance, and employee well-being, we now investigate the underlying mechanisms in more detail. As illustrated in Figure 1, the management of skills involves actions from the employee and the respective supervisor. Before jointly discussing the allocation of training, employees need to take part in the skill assessment and managers need to schedule development meetings on a regular basis. Following this process, we first consider the employee perspective, where we analyze treatment effects on participation and performance in the skill assessments. Then, we quantify the effects on supervisor behavior where we investigate the role of managerial attention in the skill management process.

5.1 Employee Behavior

In this section, we study the effect of the treatment on (i) employees' decision to take part in a skill assessment and (ii) outcomes of those who participated in the assessments. We focus on the three most important skill assessments that correspond to the majority of work assignments of the firm and for which both pre- and post-experimental data is available.[13] Employees are only invited to a skill assessment if they have been assigned to the corresponding skill. The key outcome variable that we consider here is thus average participation for assessments associated with these three skills. Our final dataset includes 19,447 completed skill assessments. Prior to the experiment, the average share of correctly answered questions was 53.03% in the most wide-spread skill assessment. In the second and third most wide-spread skill assessments, this share was 64.02%, and 71.64%, respectively. For the analyses, we standardize the results for each assessment separately and then create the mean standardized score across all three assessments.

To assess the effect of the intervention we regress assessment participation and the standardized mean assessment score on a treatment dummy. The results are reported in Table 5. In columns (2) and (4), we control for prior assessment participation and, for those who participated, their standardized test scores. As column (1) shows, the treatment reduced the average participation rate by 6.59 percentage points from 79.03% to 72.43% ($p = 0.004$).[14] Hence, rather than encouraging employees to take part in the assessment to learn about their own skill gaps when they can do so in private, the intervention reduced the likelihood that employees assess their own skills. This underscores the importance of active supervisor monitoring to ensure that employees take part in skill assessments. We find no evidence, however, that the treatment affected the average assessment result of those who took part in the skill assessment (columns 3 and 4). The coefficients for prior participation and prior scores reported in columns (2) and (4) further show that assessment participation is strongly correlated with assessment

[13]Figure A3 we show that the three basic skills cover 89% of all work assignments. In Table A2, we show treatment effects for all skill assessments in the treatment period across all six skills.
While it is possible to re-take the assessment, we find only six instances in our data set where an employee participates twice in the same assessment. Therefore, we only consider test outcomes in the first try for every assessment.

[14]In Table A5, we analyze whether the treatment has reduced the employees' reasoning about the priority of the skill assessment. We find no evidence that the treatment has lowered the perceived priority of the assessment. Therefore, employees are still aware that they are expected to complete the skill assessment, but rather consciously decided to avoid it now that non-participation is not observable for the manager.

Table 5: Treatment Effects on Assessment Participation and Assessment Scores

	Assessment Participation		Assessment Scores	
	(1)	(2)	(3)	(4)
Treatment	-0.066^{***}	-0.041^{**}	-0.015	-0.016
	(0.022)	(0.020)	(0.050)	(0.044)
Prior participation		0.375^{***}		0.341^{***}
		(0.032)		(0.073)
Prior part. × prior score		0.034^{***}		0.597^{***}
		(0.008)		(0.025)
Constant	0.790^{***}	0.475^{***}	-0.054	-0.313^{***}
	(0.015)	(0.032)	(0.034)	(0.066)
Observations	2321	2232	2210	1988

Note: Table shows estimates from linear regressions on assessment participation (columns 1 and 2) and assessment scores (columns 3 and 4). Participation denotes the number of assignments an employee has taken part in during the treatment period, devided by the number of assignments they have been assigned to. Assessment scores depict the average standardized ratio of correctly answered questions on the knowledge assessment. Prior participation captures the participation rate prior to the experiment. Prior participation × prior score is a multiplicative term of the participation rate prior to the experiment and the average standardized ratio of correctly answered questions in all skill assessments prior to the experiment. Robust standard errors clustered on the team level in parentheses, **, *** indicate significance on the 5% and 1% level, respectively.

outcomes: Employees with higher past assessment scores are more likely to participate now (column 2), and those who participated previously have substantially higher current scores compared to first-time participants (column 4).

A key conjecture at the outset of the experiment was that reducing managers' ability to observe assessment outcomes affects employees differently depending on their prior skill level. If employees with lower skill levels are discouraged from performing assessments due to fear of negative judgement, the treatment may have positively impacted participation by ensuring results remain private.[15] To investigate this, we classify those employees that had taken part in at least one of the focal skill assessments prior to the introduction of the treatment and achieved an above-median score as *high-skill* employees and the remainder as *low-skill*.[16]

[15] Our pre-registered hypotheses here were that the treatment (i) reduces test participation and test performance of highly skilled employees but (ii) increases test participation for employees with lower skills.

[16] Note that prior non-participants are thus also classified as *low-skill*. As shown in Table 5, there is a strong

We then regress assessment participation and assessment scores on a treatment indicator, interacted with a binary indicator for the skill level of an employee. The respective results are shown in Table 6. Importantly, the reduction in assessment participation seems to be driven by low-skill employees. (column 2). Hence, the intervention did *not* encourage this group to take part in the assessment. Quite to the contrary, the decrease in assessment participation is driven by this group. On the other hand, assessment participation among the *high-skill* employees is hardly reduced ($p = 0.771$). Interestingly, we find some (though statistically weak) evidence that the treatment had a stronger negative effect on assessment scores for *high-skill* employees.[17]

Hence, rather than encouraging employees with lower skill levels to assess their own skills by reducing the fear of being blamed for bad outcomes, the treatment even reduced assessment participation particularly in this group. Assessment scores are somewhat reduced for the highly-skilled employees – likely because they can no longer impress their supervisors through good test scores. It has often been claimed that reputational concerns, or as put by Holmström (2017) in his nobel lecture "the craving for appreciation and the desire to impress superiors" (p.1772), are a key driver of motivation in organizations. To study the role of such reputational concerns in more detail we turn to the post-experimental survey where we included two items to assess the role of reputational concerns for taking part in the test ("*I participate to demonstrate to others that I am competent*.") or not doing so ("*If I don't participate, it's because I expect to perform poorly.*"). First of all, the motive to participate to signal competence to the supervisor is substantially stronger than the motivation not to participate to avoid a poor performance. On a scale from 1 ("Fully Disagree") to 7 ("Fully agree") the mean is 4.71 for the

correlation between test participation and test outcomes, i.e. past non-participants have substantially lower test outcomes when taking part in a test now and current test participation is higher for employees with better past test outcomes. Additionally, Figure A5 shows that prior non-participants have less skills on average compared to any other group.

[17]We also look at more specific subgroups comparing test participation and test scores between treatment and control group for the prior non-participants and by quartile of prior assessment scores. The results are displayed in Figure A4 in the Appendix. Notably, the reduction in test participation is driven by those employees who had not taken part in prior tests.

Table 6: Treatment Effects on Assessment Outcomes by Skill Level

	Participation		Scores	
	(1)	(2)	(3)	(4)
Treatment	-0.046^{**}	-0.093^{***}	-0.022	0.044
	(0.022)	(0.028)	(0.047)	(0.063)
Treatment × High-Skill		0.099^{***}		-0.134^{*}
		(0.027)		(0.073)
High-Skill	0.107^{***}	0.058^{***}	0.772^{***}	0.837^{***}
	(0.014)	(0.017)	(0.037)	(0.052)
Constant	0.738^{***}	0.761^{***}	-0.409^{***}	-0.440^{***}
	(0.017)	(0.018)	(0.037)	(0.042)
p-value High-Skill		0.771		0.109
Observations	2232	2232	1988	1988

Note: Table shows estimates from linear regressions on assessment participation (columns 1 and 2) and assessment scores (columns 3 and 4). Participation denotes the number of assignments an employee has taken part in during the treatment period, devided by the number of assignments they have been assigned to. Assessment scores depict the average standardized ratio of correctly answered questions on the knowledge assessment. The variable 'High-Skill' takes value 1 for all employees with an above median average assessment score prior to the experiment and value 0 otherwise. 'p-value High-Skill' reports the p-value of a Wald test of Treatment + Treatment × High-Skill'. Robust standard errors clustered on the team level in parentheses, *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

former and 2.19 for the latter. To see how the treatments affect these motives we regress the response to these items on the treatment (columns 1 and 3) and the treatment interacted with the *high-skill* dummy (columns 2 and 4). As column (1) in Table 7 shows, participants in the treatment group are significantly less motivated to participate in the knowledge assessment to demonstrate their ability to their supervisor. As column (2) shows that this effect is indeed driven by high-performing employees. In contrast to our prior expectations but in line with our observations regarding the treatment effect on assessment participation, the treatment does not reduce the (already low) stated motive not to take part in test because of being afraid to do poorly (column 3 in Table 7). And while low-skill employees indeed state significantly more often that they avoid the assessment due to being afraid to do poorly than high-skill employees (columns 3 and 4), there is no evidence for differential treatment effect on this motive.

Table 7: Signaling and reputational concerns

	Demonstrate Competence		Fear of Incompetence	
	(1)	(2)	(3)	(4)
Treatment	−0.435***	−0.008	0.126	0.159
	(0.163)	(0.210)	(0.111)	(0.186)
Treatment × High-Skill		−0.796***		−0.062
		(0.292)		(0.239)
High-Skill	−0.099	0.267	−0.692***	−0.664***
	(0.146)	(0.181)	(0.119)	(0.155)
Constant	4.968***	4.773***	2.495***	2.480***
	(0.128)	(0.141)	(0.096)	(0.113)
p-value High-Skill		0.000		0.491
Observations	785	785	782	782

Note: Table shows estimates from linear regressions on the strength of an employee's competence motif when considering to participate in the skill assessment (columns 1 and 2) and the strength of an employee's avoidance motif when considering not to participate in the skill assessment (columns 3 and 4). All questions are answered on a 7-point likert scale from 1 = 'strongly disagree' to 7 = 'strongly agree'. The variable 'High-Skill' takes value 1 for all employees with an above median average assessment score prior to the experiment and value 0 otherwise. 'p-value High-Skill' reports the p-value of a Wald test of Treatment + Treatment × High-Skill'. Robust standard errors clustered on the team level in parentheses, *** indicates significance on the 1% level.

5.2 Supervisor Behavior

In this section, we study the role of managers in the skill management process. To assess the treatment effect on managerial attention, we first investigate how the treatment affected the likelihood that managers conducted developmental meetings and at which frequency. We begin by regressing the meeting incidence (as reported by employees in the post-experimental survey) on a treatment indicator. As the results reported in Table 8 show, the treatment reduced the likelihood that at least one meeting is conducted from 0.77 to 0.67, i.e. by about 13.63% (see column 1). We replicate these findings using team level data in columns (2) and (4) where we can control for the prior number of meetings, as reported in the pre-experimental survey. Hence, the intervention not only reduces employees' willingness to assess their own skill gaps but also managers' own time investments into employee development.[18] To understand the behavioral drivers of this effect it is informative to note that the treatment primarily affects the likelihood of conducting at least one meeting. As documented in columns (3) and (4), the effect on the number of meetings has a similar magnitude which shows that it is the incidence rather than the frequency that differs. In other words, supervisors do not substantially reduce the meeting frequency *after* having held at least one meeting without having access to assessment outcomes.[19] This suggests that managers are *not* reducing the frequency of meetings because they have had worse experiences as a result of the treatment. Instead, it appears that supervisors ex-ante perceive the meetings as being less valuable when they know that they no longer have access to information on employee skills. This further underlines the importance of hard information on skills levels, not only for employees, but also for managers.

In a next step we study how the reduction in meeting incidence affected employees of different prior skill levels. In the regressions reported in Table 9 we regress meeting incidence

[18]In Section A.7 we show suggestive evidence that the reduction in managerial attention also caused the decline in employee satisfaction reported in Section 4.3.

[19]To corroborate this we also can study the probability of having more than one meeting conditional on having had at least one. This is indeed essentially unaffected by the treatment: In the sample of employees who had at least one meeting the the estimate of the treatment coefficient is close to zero.

Table 8: Treatment effects on meetings

	Prob. of meeting		Num. of meetings	
	(1)	(2)	(3)	(4)
Treatment	−0.106**	−0.131***	−0.148*	−0.160*
	(0.043)	(0.043)	(0.084)	(0.089)
Prior meetings		0.231***		0.444***
		(0.057)		(0.115)
Constant	0.774***	0.539***	1.189***	0.812***
	(0.025)	(0.059)	(0.054)	(0.113)
Level of Analysis	Individual	Team	Individual	Team
Observations	846	95	846	95

Note: Table shows estimates from linear regressions of the likelihood or frequency of development meetings on a treatment indicator. For the individual-level analyses, the dependent variables are a binary indicator that takes value 1 if an employee has indicated to have had more than one meeting with their supervisor (1) and the overall number of meetings (3), which can take values between 0 and 3. For the team-level analysis, the dependent variables are the number of employees with at least one meeting per team, relative to the total number of employees per team who responded to the post-experimental survey (2) and the average number of meetings summarized across each team (4). Columns (2) and (4) further include the self-reported number of meetings per team as control variable, which was assessed in the pre-experimental survey. Not every team met the mininum participation threshold of 5 responses, which is why several teams could not be included in this analysis. Re-running the team-level analyses without control variables does not change the results. Robust standard errors (clustered on the team level in columns 1 and 3) in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

on the treatment now again for high- and low-skill employees separately. Interestingly, we find that the likelihood of having at least one development meeting is, in particular, substantially reduced among high-skill employees by 18.48% (p = 0.012). For low-skill employees, the negative treatment effect is less pronounced at 6.87% and not statistically significant.[20] Apparently, as high-skill employees no longer can demonstrate their abilities (as shown in Table 7) the managers also find it less rewarding to meet with them, for instance as it is harder to provide tangible recognition and have a fruitful discussion on their skill set.[21] The reduction in meeting frequency among high-skill employees may appear less concerning at first glance when in particular the low-skilled employees benefit more from more managerial attention. However, it is important to recall that as laid out in section Section 2, a key aim of the firm is not only the preserve the skills for routine tasks, but also to broader the skill set of employees to adapt to technological changes. Indeed employees classified as high-skilled tend tend to have a larger number of skills that are assigned to them (see Figure A5 in the Appendix). The substantial reduction in meetings thus may also be problematic in this group. We will investigate this in more detail in the next section.

[20]This further shows that individual participation in the skill assessment does not appear to influence the likelihood of arranging development meetings, as the reduction in assessment participation is driven by *low-skill* employees (see Table 6). If assessment participation and managerial attention were directly correlated, we would likely observe the opposite effect in Table 9. This is in line with our interpretation that managers make decisions on the allocation of managerial attention directly in response to the treatment.

[21]Note that the patterns reported in Table 7 are not driven by the reduction in meetings for high performers. The results remain virtually unchanged when controlling for meeting incidence or frequency as shown in Table A7.

Table 9: Treatment effects on meetings by skill level

	Prob. of meeting	
	High-Skill	Low-Skill
Treatment	-0.144^{**}	-0.052
	(0.057)	(0.052)
Constant	0.778^{***}	0.761^{***}
	(0.034)	(0.032)
Observations	419	362

Note. Table shows estimates from linear regressions of the likelihood of a development meetings on a treatment indicator, We split the sample based on the variable 'High-Skill' which takes value 1 for all employees with an above median average assessment score prior to the experiment (column 'High-Skill') and value 0 otherwise (column 'Low-Skill'). Robust standard errors clustered on team in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

6 Training Assignment and Work Performance

In the previous sections we presented evidence showing that the treatment reduced skill assessment participation among low-skill employees, but reduced managerial attention in particular for the high-skilled. In the next step, we explore to what extent this heterogeneity influences the training intensity and type of training received by employees in these two groups, and how this, in turn, can help in understanding the performance effects of skill management. As Table 10 shows, the treatment reduced the training intensity in both groups. While there is no significant effect heterogeneity (see Table A9), the reduction appears particularly pronounced among high-skill employees. This naturally raises the question of whether the reduction in this group is concerning, given that skill levels are already high at the outset. To this end it is important to consider different purposes of trainings. As laid out in section Section 2 the firm distinguishes between trainings to fill gaps in skills employees are already supposed to possess to raise their performance on standard tasks and trainings to acquire new skills – for instance learning to work on tasks based on new technologies that will become more important in the future or are needed to solve more complex and non-routine problems. When the reduction of trainings for the high-skill employees is concentrated on the former, indeed it may not be very detrimental or could even increase efficiency eliminating useless time expenditures for trainings. If, however, the treatment affects the latter form of trainings it can be harmful undercutting firms' abilities to adapt and prepare their high performing employees for changing skill requirements.

Table 10: Treatment Effects by Employee Type

	Num. of Trainings	
	High-Skill	Low-Skill
Treatment	−0.272**	−0.171*
	(0.116)	(0.092)
Constant	0.463***	0.348***
	(0.096)	(0.077)
Observations	1043	1191

Note: Table shows estimates from linear regressions on the number of trainings booked by prior skill-level. We split the sample based on the variable 'High-Skill' which takes value 1 for all employees with an above median average assessment score prior to the experiment (column 'High-Skill') and value 0 otherwise (column 'Low-Skill'). Robust standard errors clustered on the team level in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

When managers register an employee for a training, they have to indicate whether the purpose of the training is to maintain an existing skill (referred to as *skill preservation*) or to develop a new skill (*new skill acquisition*).[22] We can make use of this distinction to investigate which types of trainings are particularly affected in which employee group.[23] Figure 3 shows the average number of trainings, split by training type (skill preservation, new skill acquisition) and employee skill level (high, low). The corresponding regression results are reported in Table A10 in the Appendix. While training intensity is lower across both groups and training types, the effect is particularly pronounced for trainings dedicated the acquisition of new skills among high-skill employees ($p = 0.01$, panel C) and (to a lesser extent) for skill preservation trainings for the low-skilled ($p = 0.066$, panel B).

[22]Supervisors can also select the option 'not relevant to a specific skill'. However, this option is rarely selected, so we drop this alternative from the analysis.

[23]In Table A8 in the Appendix, we investigate whether the overall number of trainings dedicated to skill preservation and acquisition is differentially effected by the treatment. We do not find such an effect. If anything, the treatment effect is slightly larger for training dedicated to the acquisition of new skills.

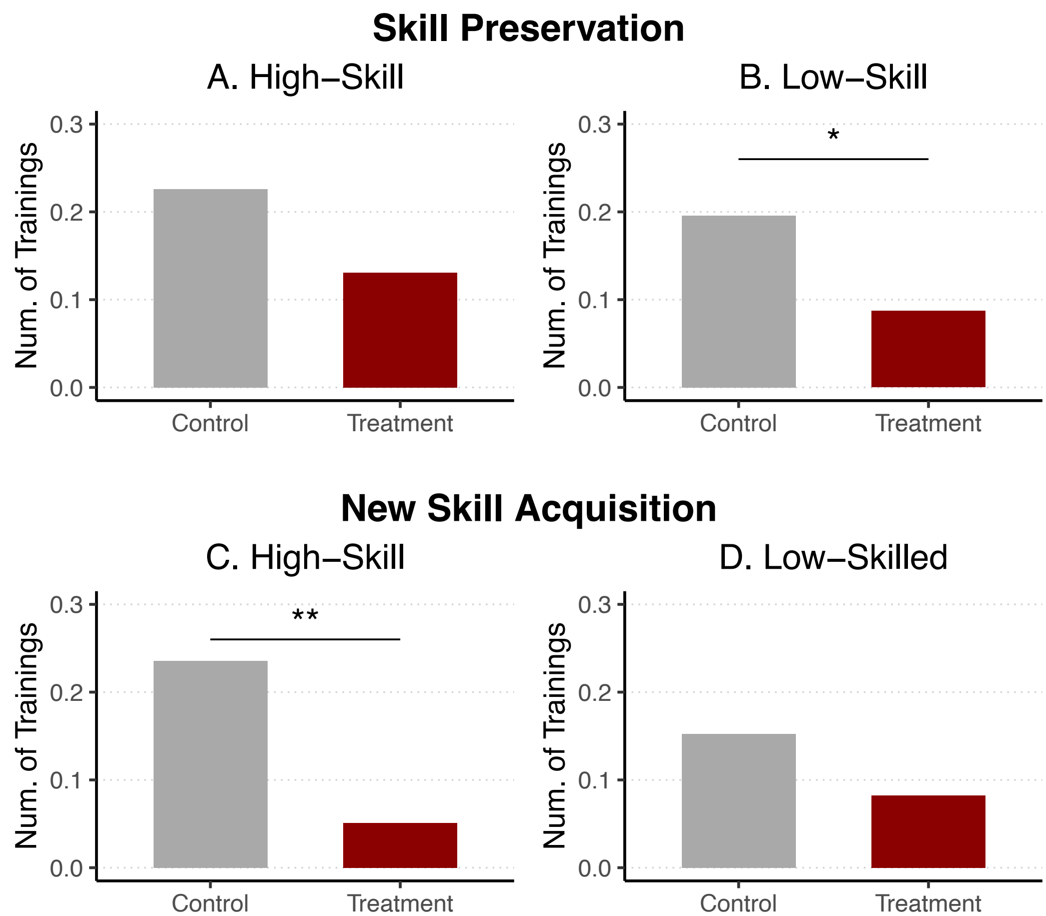

Figure 3: Treatment effects by Employee Type and Training Purpose

The observation that the reduction in trainings is particularly pronounced for new skill acquisition of high-skill employees may suggest that the overall reduction in work performance is concentrated on work assignments that require a higher level of skill or a broader skill set in general. To investigate this conjecture, we study heterogenous treatment effects by the degree of assignment complexity. According to the firm's partition of work assignments into 13 different types, each type can be classified as either *routine* or *complex*. Routine assignments often include solving problems that occur on a regular basis or products that require little technical sophistication to work with. Complex assignments often require non-standard custom solutions that have a higher risk of failure. These assignments are either associated with recent technologies that are inherently complex or a sophisticated problem that other technicians were previously unable to resolve. Therefore, assignments can be re-classified from routine to complex. In line with the firm's description, our final classification reveals that routine assignments are performed more regularly and take relatively little time (M = 14.5 min.), while complex assignments take more than twice as long (M = 32.8 min.), and occur less frequently (see Section A.12). In Table 11, we run separate difference-in-difference regressions for tasks associated with either routine or complex work assignments with the average execution times as the dependent variable. As column (1) shows, there is no discernible treatment effect on the weekly average completion time for routine assignments. But we indeed find a 3.169 minute increase in the weekly average completion time for complex tasks (column 2), which corresponds to a 9.7% increase relative to the post-treatment average in the control group. In summary, we find that the reduction in training intensity is particularly strong for new skill acquisition among high-skill employees such that the reduction in active skill management in particular leads to a decline in upskilling among higher performing employees. In line with this pattern, the reduction in work performance is concentrated in complex work assignments.

Table 11: Treatment effects on Work Performance by Assignment Type

	Routine	Complex
Treatment	-0.044	3.169**
	(0.184)	(1.568)
Team FE	Yes	Yes
Task FE	Yes	Yes
Product × Week FE	Yes	Yes
Mean Dep. Var	14.53	32.84
Observations	128 146	21 266

Note: Table shows average treatment effects from difference-in-difference regressions on the average time per task in minutes. Column (1) shows results for work assignments classified as routine assignments, whereas column (2) shows results for complex assignments. For all regressions, we focus on the four key tasks used in the previous analysis. All models include team, task, as well as Product × Week fixed effects. Averages from the control group are depicted below the estimates.Standard errors clustered on the team-level in parentheses. ** indicates significance on the 5% level.

7 Conclusion

We studied the value of skill management in organizations through a natural field experiment. Our findings highlight that active skill management is crucial for human capital investments and firm performance. Even a slight reduction in managers' ability to monitor and manage employee skills significantly decreased training intensity and productivity. Additionally, contrary to existing literature in psychology and behavioral economics, we observed that reduced supervisor monitoring also led to a decline in employee well-being.

We further studied the mechanisms of skill management. Reducing managers' ability to monitor employee skills diminished employees' willingness to assess their own skill gaps, particularly among low skill employees. Instead of encouraging these employees to identify training needs without fear of blame, the removal of monitoring had the opposite effect. High-skill employees, on the other hand, did not reduce their participation but appeared to reduce their effort in these assessments. Our post-experimental survey indicates that this behavior is driven by reduced reputational incentives when employees know their assessment outcomes are not observed. Both these findings highlight the significance of moral hazard issues in training investments. Despite the benefits of human capital gains for both groups of employees, there is thus a tendency to under-invest and active skill management helps to mitigate this problem. Granting supervisors access to hard information about employee skills commits employees to invest in their own learning and fosters managers' time investments into employee development. Removing access to this information reduces the likelihood and frequency of development meetings, especially for high-skill employees. Having documented the main effects and general mechanics of skill management, we set out to understand the connection between the assignment of training and team productivity in more detail. Sandvik et al. (2024), for instance, found that uptake of voluntary training is lower among low-performers, even though they benefit from training the most. In environments undergoing technological changes it is, however, less clear whether prioritizing the skill gaps of low-performing employees should take precedence over broadening the skill set of higher performing employees, as firms constantly need to adapt to changing skill requirements. Our analysis indeed shows that less extensive skill management induced by the treatment particularly reduced managerial attention and subsequent training intensity among high-skill employees. Moreover, this reduction in training intensity predominantly affected trainings intended to acquire new skills. In turn, the intervention led to a reduction in work performance in complex, non-routine tasks. These findings add to the literature on heterogeneity in training goals across industries (Alfonsi et al., 2020; Caicedo, Espinosa, & Seibold, 2022). In workplace environments with heavily standardized workflows, closing skill gaps among low performers should take priority, as illustrated by a recent study from Adhvaryu, Nyshadham, & Tamayo (2022), who highlight the value of refresher trainings in a fast food chain. Conversely, jobs subject to frequent technological change and innovation may benefit more from further developing the skill set of high-performing employees. In summary, our results demonstrate that active skill management and close managerial scrutiny of skill levels are crucial for organizational performance. It fosters the willingness of low-skilled employees to assess their own skills, encourages high-skill employees to improve and acquire new skills, and ensures that leaders invest time in helping both groups strengthen their human capital.

8 References

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043–1171). Elsevier. https://doi.org/10.1016/S0169-7218(11)02410-5

Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in US wage inequality. *Econometrica*, *90*(5), 1973–2016. https://doi.org/10.3982/ECTA19815

Adhvaryu, A., Kala, N., & Nyshadham, A. (2022). Management and Shocks to Worker Productivity. *Journal of Political Economy*, *130*(1), 1–47. https://doi.org/10.1086/717046

Adhvaryu, A., Nyshadham, A., & Tamayo, J. (2022). *An anatomy of performance monitoring*. Harvard Business School Working Paper.

Adhvaryu, A., Nyshadham, A., & Tamayo, J. (2023). Managerial quality and productivity dynamics. *The Review of Economic Studies*, *90*(4), 1569–1607. https://doi.org/10.1093/restud/rdac062

Alfonsi, L., Bandiera, O., Bassi, V., Burgess, R., Rasul, I., Sulaiman, M., & Vitali, A. (2020). Tackling youth unemployment: Evidence from a labor market experiment in uganda. *Econometrica*, *88*(6), 2369–2414. https://doi.org/10.3982/ECTA15959

Alfonsi, L., Bassi, V., Rasul, I., & Spadini, E. (2024). *The returns to skills during the pandemic: Experimental evidence from uganda* (No. 32785). National Bureau of Economic Research. https://doi.org/10.3386/w32785

Autor, D., Chin, C., Salomons, A., & Seegmiller, B. (2024). New frontiers: The origins and content of new work, 1940–2018. *The Quarterly Journal of Economics*, 1399–1465. https://doi.org/10.1093/qje/qjae008

Bapna, R., Langer, N., Mehra, A., Gopal, R., & Gupta, A. (2013). Human capital investments and employee performance: An analysis of IT services industry. *Management Science*, *59*(3), 641–658. https://doi.org/10.1287/mnsc.1120.1586

Bartel, A. P. (1989). *Formal employee training programs and their impact on labor productivity: Evi-*

dence from a human resources survey. https://doi.org/10.3386/w3026

Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, *96*(5), 1652–1678. https://doi.org/10.1257/aer.96.5.1652

Berge, L. (2018). *Efficient estimation of maximum likelihood models with multiple fixed-effects: The r package FENmlm*.

Black, S. E., & Lynch, L. M. (1996). Human-capital investments and productivity. *American Economic Review*, *86*(2), 263–267. https://doi.org/http://www.jstor.org/stable/2118134

Bloom, N., Sadun, R., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, *122*(4), 1351–1408. https://doi.org/10.1162/qjec.2007.122.4.1351

Bloom, N., Sadun, R., & Van Reenen, J. (2010). Recent advances in the empirics of organizational economics. *Annual Review of Economics*, *2*(1), 105–137. https://doi.org/10.1146/annurev.economics.050708.143328

Bloom, N., Sadun, R., & Van Reenen, J. (2012). The organization of firms across countries*. *The Quarterly Journal of Economics*, *127*(4), 1663–1705. https://doi.org/10.1093/qje/qje029

Brunello, G., & Wruuck, P. (2021). Skill shortages and skill mismatch: A review of the literature. *Journal of Economic Surveys*, *35*(4), 1145–1167. https://doi.org/10.1111/joes.12424

Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machines learn and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, *108*, 43–47. https://doi.org/10.1257/pandp.20181019

Buell, R., Cai, W., & Sandino, T. (2022). *Learning or playing? The effect of gamified training on performance*. https://doi.org/10.2139/ssrn.4307211

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2013). Overconfidence and Social Signalling. *The Review of Economic Studies*, *80*(3), 949–983. https://doi.org/10.1093/restud/rds046

Cai, W., Chen, Y., & Chen, Y. (2024). *Empowered learning as a control system*.

Caicedo, S., Espinosa, M., & Seibold, A. (2022). Unwilling to train?—firm responses to the colombian apprenticeship regulation. *Econometrica*, *90*(2), 507–550. https://doi.org/10.3982/ECTA18911

Campbell, D., Epstein, M. J., & Martinez-Jerez, F. A. (2011). The learning effects of monitoring. *The Accounting Review*, *86*(6), 1909–1934. https://doi.org/10.2308/accr-10129

Christ, M. H. (2013). An experimental investigation of the interactions among intentions, reciprocity, and control. *Journal of Management Accounting Research*, *25*(1), 169–197. https://doi.org/10.2308/jmar-50443

De Grip, A., & Sauermann, J. (2012). The effects of training on own and co-worker productivity: Evidence from a field experiment. *The Economic Journal*, *122*(560), 376–399. https://doi.org/10.1111/j.1468-0297.2012.02500.x

De Janvry, A., He, G., Sadoulet, E., Wang, S., & Zhang, Q. (2023). Subjective performance evaluation, influence activities, and bureaucratic work behavior: Evidence from china. *American Economic Review*, *113*(3), 766–799. https://doi.org/10.1257/aer.20211207

Deci, E. L., Connell, J. P., & Ryan, R. M. (1989). Self-determination in a work organization. *Journal of Applied Psychology*, *74*(4), 580. Retrieved from https://doi.org/10.1037/0021-9010.74.4.580

Delfino, A., Garnero, A., Inferrera, S., & Sadun, R. (2024). *Unwilling to reskill? Evidence from a survey experiment with jobseekers*.

Diaz, B., Neyra-Nazarrett, A., Ramirez, J., Sadun, R., & Tamayo, J. (2024). *Training within firms*.

Espinosa, M., & Stanton, C. T. (2023). *Training, communications patterns, and spillovers inside organizations*. https://doi.org/10.3386/w30224

Ewers, M., & Zimmermann, F. (2015). Image and Misreporting. *Journal of the European Economic Association*, *13*(2), 363–380. https://doi.org/10.1111/jeea.12128

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, *96*(5), 1611–1630. Retrieved from https://doi.org/10.1257/aer.96.5.1611

Friebel, G., Heinz, M., & Zubanov, N. (2022). Middle managers, personnel turnover, and performance: A long-term field experiment in a retail chain. *Management Science*, *68*(1), 211–229. https://doi.org/10.1287/mnsc.2020.3905

Gosnell, G. K., List, J. A., & Metcalfe, R. D. (2020). The impact of management practices on employee productivity: A field experiment with airline captains. *Journal of Political Economy*, *128*(4), 1195–1233.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, *73*, 103–130. https://doi.org/10.1016/j.euroecorev.2014.10.006

Hoffman, M., & Stanton, C. T. (2024). *People, practices, and productivity: A review of new advances in personnel economics*. National Bureau of Economic Research. https://doi.org/10.3386/w32849

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, *10*(1), 74. https://doi.org/10.2307/3003320

Holmström, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, *13*(2), 324. https://doi.org/10.2307/3003457

Holmström, B. (2017). Pay for performance and beyond. *American Economic Review*, *107*(7), 1753–1777. https://doi.org/10.1257/aer.107.7.1753

Jackson, C. K., & Schneider, H. S. (2015). Checklists and Worker Behavior: A Field Experiment. *American Economic Journal: Applied Economics*, *7*(4), 136–168. https://doi.org/10.1257/app.20140044

Jensen, N., Lyons, E., Chebelyon, E., Bras, R. L., & Gomes, C. (2020). Conspicuous monitoring and remote work. *Journal of Economic Behavior & Organization*, *176*, 489–511. https://doi.org/10.1016/j.jebo.2020.05.010

Kantor, J., & Sundaram, A. (2022, August 14). The rise of the worker productivity score. The new york times. Retrieved from https://www.nytimes.com/interactive/2022/08/

14/business/worker-productivity-tracking.html

Katok, E., & Siemsen, E. (2011). Why genius leads to adversity: Experimental evidence on the reputational effects of task difficulty choices. *Management Science*, *57*(6), 1042–1054. https://doi.org/10.1287/mnsc.1110.1331

Manthei, K., & Sliwka, D. (2019). Multitasking and subjective performance evaluations: Theory and evidence from a field experiment in a bank. *Management Science*, *65*(12), 5861–5883. https://doi.org/10.1287/mnsc.2018.3206

Manthei, K., Sliwka, D., & Vogelsang, T. (2023). Talking about performance or paying for it? A field experiment on performance reviews and incentives. *Management Science*, *69*(4), 2198–2216. https://doi.org/10.1287/mnsc.2022.4431

Merchant, K., & Van der Stede, W. (2017). *Management control systems - performance measurement, evaluation and incentives*. London, UK: Pearson.

Mincer, J. (1974). The human capital earnings function. In *Schooling, experience, and earnings* (pp. 83–96). NBER. Retrieved from http://www.nber.org/books/minc74-1

Murty, R. N., & Karanth, S. (2022, October 27). Monitoring individual employees isn't the way to boost productivity. Harvard business review. Retrieved from https://hbr.org/2022/10/monitoring-individual-employees-isnt-the-way-to-boost-productivity

Nebeker, D. M., & Tatum, B. C. (1993). The effects of computer monitoring, standards, and rewards on work performance, job satisfaction, and stress. *Journal of Applied Social Psychology*, *23*(7), 508–536. https://doi.org/10.1111/j.1559-1816.1993.tb01101.x

OECD. (2017). *Getting skills right: Skills for jobs indicators*. OECD. https://doi.org/10.1787/9789264277878-en

Ravid, D. M., White, J. C., Tomczak, D. L., Miles, A. F., & Behrend, T. S. (2023). A meta-analysis of the effects of electronic performance monitoring on work outcomes. *Personnel Psychology*, *76*(1), 5–40.

Rochambeau, G. de. (2019). *Monitoring and intrinsic motivation: Evidence from liberia's trucking* *firms*.

Sandvik, J., Saouma, R. E., Seegert, N., & Stanton, C. (2024). Can firms use self-selection to improve the efficacy of human capital investments? Evidence from a field experiment. *Evidence from a Field Experiment (July 19, 2024)*. https://doi.org/10.2139/ssrn.4899777

Sandvik, J., Saouma, R., Seegert, N., & Stanton, C. (2020). Workplace knowledge flows. *The Quarterly Journal of Economics*, *135*(3), 1635–1680. https://doi.org/10.1093/qje/qjaa013

Shrikant, A. (2023, September 8). Companies use AI to monitor workers and 45% of employees say it has a negative effect on their mental health. CNBC make-IT. Retrieved from https://www.cnbc.com/2023/09/08/employers-using-ai-to-monitor-workers-has-negative-impact-on-employees.html

Silva, J. S., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 641–658. https://doi.org/10.1162/rest.88.4.641

A Appendix

A.1 Work Assignment Process

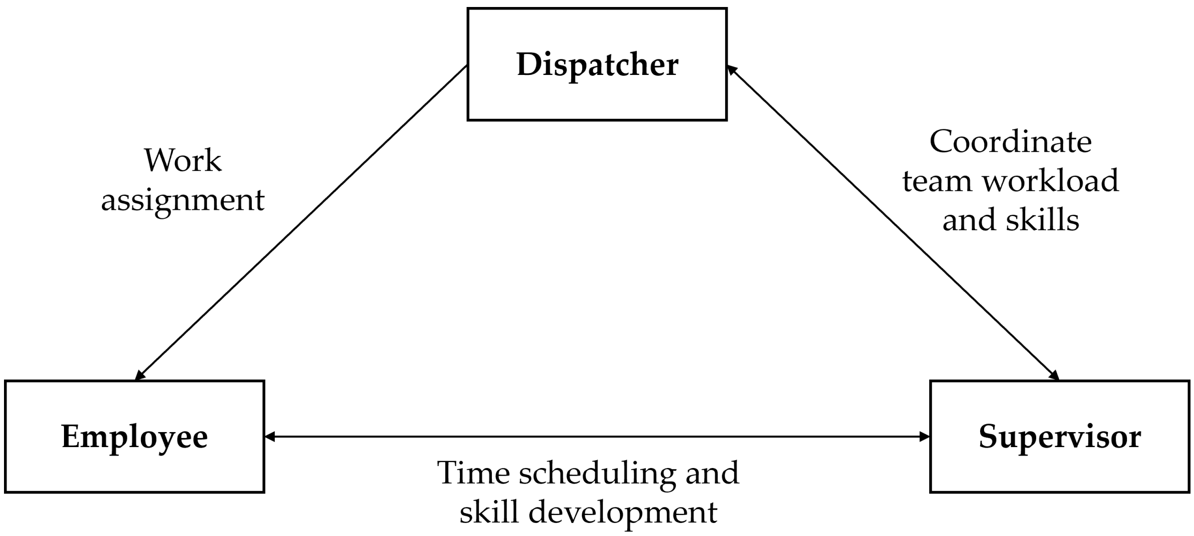

Figure A1: Distribution of Responsibilites in the Firm

A.2 Communicating the treatment

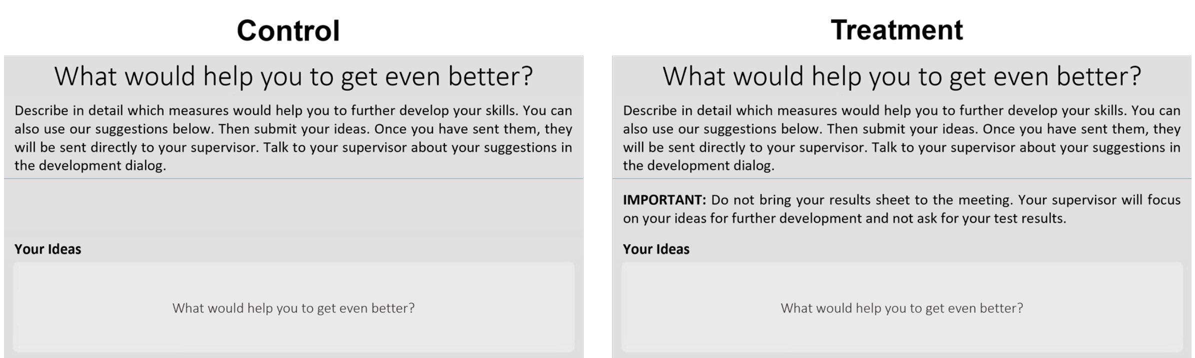

Figure A2: Communicating the treatment on the skill assessment platform

A.3 Robustness checks

Given that the number of trainings an employee receives in a given time frame can be modeled as count data, we consider alternative specifications for the main analyses conducted in Table 2. For the analysis conducted in Table A1, we consider a Pseudo-Poisson-ML estimator, which performs well on over-dispersed count data (Silva & Tenreyro, 2006). Estimations are performed in R using the *fixest* package (Berge, 2018). To interpret the results from columns (1) and (2) of Table A1, we calculate the incidence ratio by exponentiating the treatment coefficients. Overall, we find that the results are similar to those reported in the main text. For the baseline treatment comparison without covariates, the incident ratio is 0.481, which is a 48.11% reduction in the number of trainings. Including prior participation in and scores of the skill assessment changes the incident ratio to 0.479, which is equal to a 47.92% reduction in the number of trainings. In the next step, we re-analyze the treatment effect on whether or not an employee receives any training at all. Again, we find that employees in the treatment group are about 48.96% less likely to receive at least one training during the treatment period. Finally, columns (5) and (6) show team-level results, without (column 5) and including pre-treatment data (column 6). While we focus our analyses on the key skills almost all employees possess and for which we have pre and post treatment data, we further analyzed treatment effects on the remaining skill assessments that took place in the treatment period. As Table A2 shows, the reduction in assessment participation is strongest among the most wide-spread skills. In this table, the number of observations denotes all employees that were assigned to the knowledge assessment at the time it took place. As lined out previously, the number of observations is much smaller for less wide-spread skills. However, including all skills into the assessment does not change our results (see column 7)

Table A1: Treatment Effects on Training (Poisson/Logit)

	Num. of Trainings		Prob. of Training		Num. of Trainings	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	−0.732**	−0.736**	−0.714**	−0.690**	−0.636*	−0.631*
	(0.340)	(0.355)	(0.328)	(0.336)	(0.325)	(0.323)
Prior participation		1.196**		1.413***		
		(0.533)		(0.505)		
Prior trainings						0.036*
						(0.018)
Prior score		0.066		0.103		
		(0.059)		(0.075)		
Constant	−0.983***	−1.963***	−1.702***	−2.873***	2.158***	1.949***
	(0.207)	(0.526)	(0.205)	(0.516)	(0.199)	(0.241)
Level of Analysis	Individual	Individual	Individual	Individual	Team	Team
Observations	2582	2078	2582	2078	125	125
Model	Poisson	Poisson	Logit	Logit	Poisson	Poisson

Note: Table shows estimates from pseudo-poisson regressions on the the number of trainings as a dependent variable, which is the sum of all trainings associated with key skills (column (1), (2), (5) and (6). In columns (3) and (4) the dependent variable is a binary indicator for receiving training or not which takes value 1 if an employee has received at least one training and value 0 otherwise. In columns (5) and (6), we use the number of trainings on the team level as the dependent variable. Prior participation captures the participation rate prior to the experiment. Prior score denotes the average standardized ratio of correctly answered questions on a skill assessment. Prior trainings denotes the number of trainings per team prior to the experiment Robust standard errors (clustered on the team level in column 1-4) in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

Table A2: Treatment Effects on Assessment Participation (All Skills)

	Skill 1	Skill 2	Skill 3	Skill 4	Skill 5	Skill 6	All Skills
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	−0.094***	−0.046*	0.001	−0.035	−0.034	−0.147**	−0.066***
	(0.023)	(0.025)	(0.046)	(0.053)	(0.041)	(0.069)	(0.022)
Constant	0.855***	0.793***	0.649***	0.646***	0.897***	0.931***	0.783***
	(0.014)	(0.018)	(0.037)	(0.039)	(0.029)	(0.031)	(0.015)
Observations	2286	2032	1068	450	341	109	2329

Note: Tables shows estimates from linear regressions of each skill assessment during the treatment period on a treatment indicator (columns 1-6) and the average of all post-treatment assessments (column 7). Robust standard errors clustered on team in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

A.4 Participation in the Post-Experimental Survey

Table A3: Pre-treatment characteristics of survey participants

	Control (N=503)		Treatment (N=420)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p
Assessment Participation	0.88	0.24	0.87	0.27	-0.01	0.69
Assessment Score (Std.)	0.05	0.84	0.01	0.86	-0.04	0.48

Note: Table shows pre-experimental data for employees that participated in the post-experiment survey, split by treatment group. Assessment participation denotes the average participation for the tests every member of the team has been assigned to. Assessment results depict the average standardized ratio of correctly answered questions on a knowledge test. P-values are calculated using OLS, regressing the dependent variable on a treatment indicator.

A.5 Skills and Work Assignments in the Firm

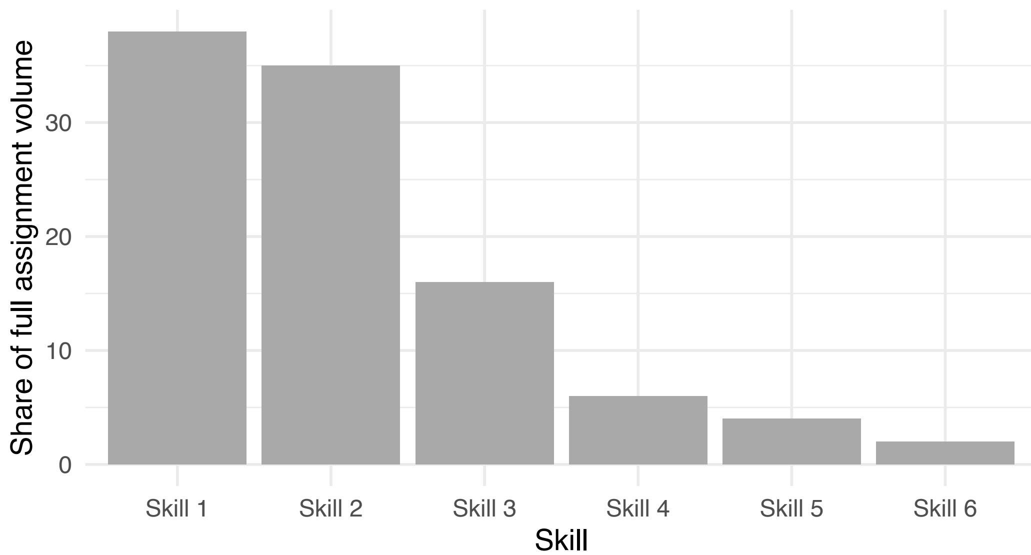

Figure A3: Relative share of work assignment volume by skill

A.6 Treatment Effects by Prior Performance

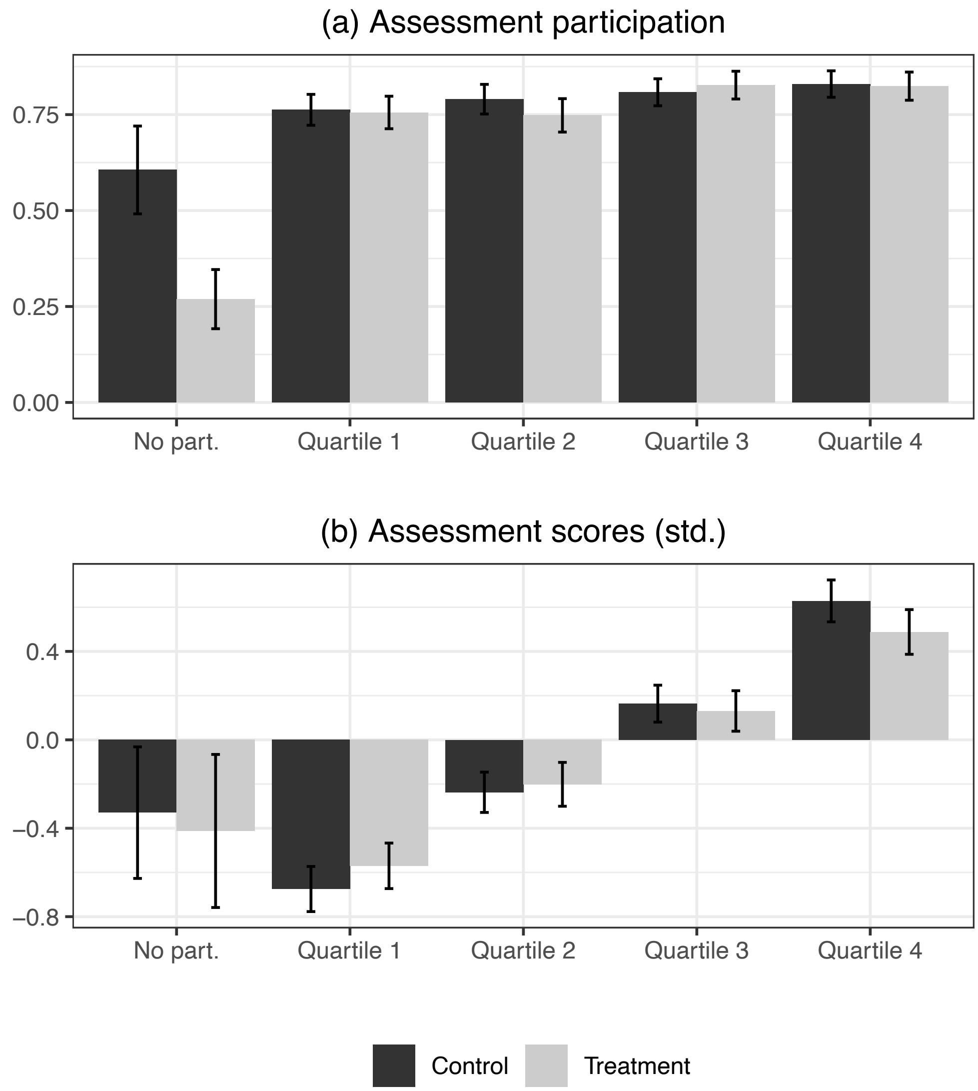

Figure A4: Outcomes by prior test participation and performance

A.7 Managerial Attention and Job Satisfaction

Our main results raise the question whether the reduction in managerial attention can explain the reduction in employee job satisfaction we documented in section Section 4.3. To study this, we regress again job satisfaction on the treatment interacted with a binary indicator that is 1 for employees who report having met at least once with their supervisor to discuss their development and 0 otherwise. While we caution that this regression has no clean causal interpretation as the meeting incidence is of course affected by the treatment, we still consider the analysis as potentially insightful. As shown in column 2 of Table A4, we find that the reduced job satisfaction is concentrated among employees who did not have any development meeting during the treatment period. We conclude that removing the managers' access to information about skill assessments leads to less supervisor interaction, which in turn appears to have reduced employee well-being.

Table A4: Managerial Attention and Job Satisfaction

	Job Satisfaction	
	(1)	(2)
Treatment	−0.200**	−0.395**
	(0.077)	(0.153)
Treatment x Meeting Conducted		0.323*
		(0.171)
Meeting Conducted		0.216
		(0.134)
Constant	0.094*	−0.074
	(0.053)	(0.124)
Observations	775	775

Note: Table shows estimates from linear regressions on job satisfaction as reported by the employees in a post-experimental survey. Job satisfaction is measured using asingle-item measure from the Linked Personnel Panel (Kampkoetter et al., 2016) and answered on a 7-point likert scale from 1 = 'strongly disagree' to 7 = 'strongly agree'. The variable 'Meeting Conducted' takes value 1 if the employee has received at least one development meeting with their supervisor and 0 otherwise. Robust standard errors clustered on the team level in parentheses. *, ** indicate significance on the 10% and 5% level, respectively.

A.8 Number of Skills by Prior Performance

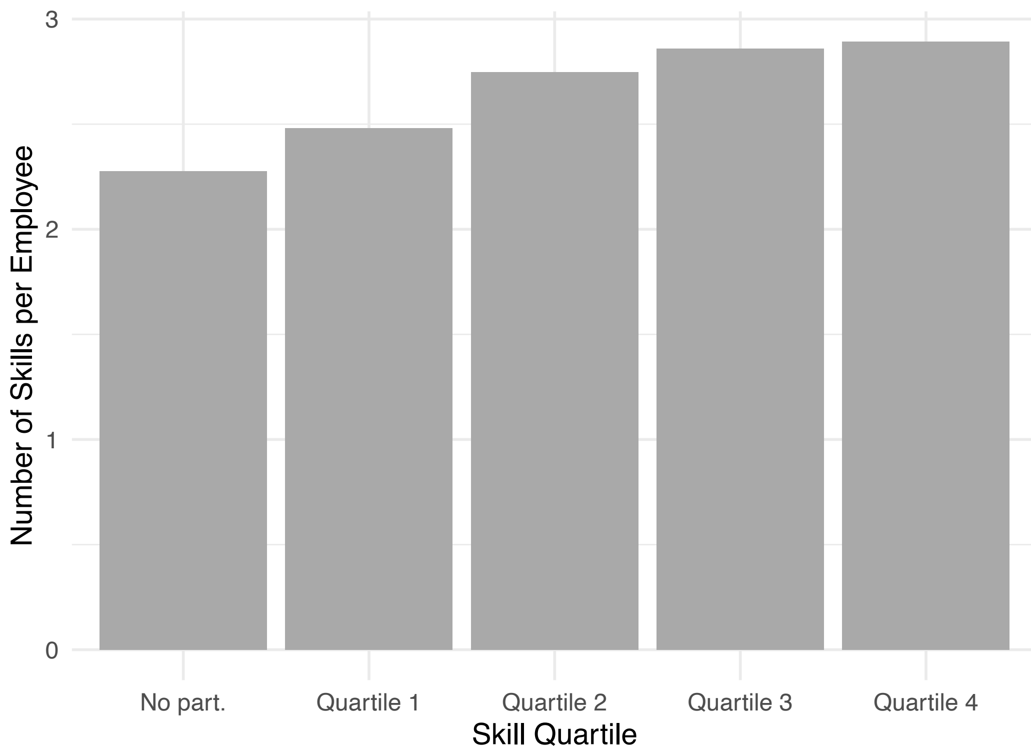

Figure A5: Average number of skills per quartile

A.9 Reasons for Participating in the Skill Assessments

Table A5: Reasons for Skill Assessment Participation

	Obligation	Demand	Improvement
	(1)	(2)	(3)
Treatment	-0.011	-0.026	-0.101
	(0.063)	(0.069)	(0.081)
Constant	0.005	0.012	0.046
	(0.044)	(0.046)	(0.052)
Observations	850	850	850

Note: Table shows results from linear regressions on questions from the post-experimental survey. Columns capture the perceived obligation to take the test ('I participate in the assessment because it is part of my job.'), supervisor demand ('I participate in the assessment because my supervisor expects me to do it.') and personal improvement ('I participate because it helps me to get better at my job.'). Robust standard errors clustered on team in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

A.10 Treatment Effects among Employees with Meetings

Table A6: Treatment effects among employees who had meetings

	Participation		Test Scores		Trainings	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	−0.065**	−0.054**	0.024	0.015	−0.121	−0.138
	(0.028)	(0.024)	(0.088)	(0.084)	(0.172)	(0.181)
Prior participation		0.331***		0.276		0.512***
		(0.067)		(0.199)		(0.164)
Prior part. × prior score		0.044***		0.638***		−0.083
		(0.016)		(0.051)		(0.052)
Constant	0.844***	0.544***	0.079	−0.195	0.466***	0.050
	(0.019)	(0.064)	(0.050)	(0.182)	(0.111)	(0.131)
Observations	577	565	570	530	614	565

Note: Table shows estimates from linear regressions on assessment participation (columns 1 and 2), assessment scores (columns 3 and 4) and the number of trainings per employee (columns 5 and 6) on a treatment indicator and further control variables. Robust standard errors clustered on the team level in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

Table A7: Signaling and reputational concerns (controlling for meeting incidence)

	Demonstrate Competence		Fear of Incompetence	
	(1)	(2)	(3)	(4)
Treatment	0.053	0.080	0.158	0.152
	(0.208)	(0.204)	(0.184)	(0.184)
Treatment × High-Skill	−0.723**	−0.847***	−0.082	−0.069
	(0.292)	(0.295)	(0.238)	(0.236)
High-Skill	0.262	0.297	−0.658***	−0.662***
	(0.186)	(0.188)	(0.154)	(0.154)
Meeting Conducted	0.956***		−0.094	
	(0.168)		(0.121)	
Number of Meetings		0.496***		−0.072
		(0.073)		(0.051)
Constant	4.038***	4.172***	2.538***	2.553***
	(0.181)	(0.156)	(0.139)	(0.118)
p-value High-Skill		0.000		0.491
Observations	781	781	778	778

Note: Column (1) and (2) show estimates from linear regressions on the strength of an employee's competence motif when considering to participate in the skill assessment. Columns (3) and (4) show estimates from linear regressions on the strength of an employee's avoidance motif when considering not to participate in the skill assessment. All questions are answered on a 7-point likert scale from 1 = 'strongly disagree' to 7 = 'strongly agree'. The variable 'High-Skill' takes value 1 for all employees with an above median average assessment score prior to the experiment and value 0 otherwise. 'p-value High-Skill' reports the p-value of a Wald test of Treatment + Treatment × High-Skill'. Robust standard errors clustered on the team level in parentheses, *** indicates significance on the 1% level.

A.11 Heterogeneous Effects by Training and Employee Type

In the following, we provide supplementary analyses for Section 6. We investigate whether the treatment deferentially affected the two types of training that are skill acquisition and skill preservation. To do so, we interact the training purpose with the number of trainings per empoyee. As columns (2) and (4) show , we find no evidence that the treatment differentially affects the types of training, as the interaction between the treatment and the purpose of training is not significant. If anything, the treatment effect on training inputs dedicated to skill acquisition is sightly larger compared to trainings dedicated to skill preservation. In Table A9, we investigate whether there are heterogenous treatment effects on traiming intensity for high-skill and low-skill employees. To test this, we regress the number of trainings on a treatment dummy interacted with the dummy for high-skill employees. While on average the reduction in training intensity is larger for high-skill employees the difference between the two groups is not statistically significant.

Table A8: Treatment Effects on Training by Type

	Num. of Trainings		Prop. of Training	
	(1)	(2)	(3)	(4)
Treatment	−0.108**	−0.122**	−0.055**	−0.059**
	(0.051)	(0.055)	(0.024)	(0.026)
Treatment × Skill pres.	0.015	0.021	0.005	0.006
	(0.065)	(0.070)	(0.027)	(0.029)
Skill preservation	0.024	0.018	0.020	0.018
	(0.054)	(0.059)	(0.021)	(0.024)
Prior participation		0.133***		0.064***
		(0.036)		(0.020)
Constant	0.175***	0.080*	0.088***	0.043*
	(0.044)	(0.045)	(0.020)	(0.024)
p-value: Skill pres.	0.121	0.112	0.091	0.089
Observations	5164	4468	5164	4468

Note: Table shows estimates from linear regressions on the number of assigned trainings per employee. In columns (1) and (2) the dependent variable is denoted by the number of trainings , while columns (3) and (4) cover the probability of training. Skill preservation (or Skill pres.) is a binary indicator that is equal to 1 if the training purpose is the preservation of an existing skill and 0 otherwise. Prior participation captures the participation rate prior to the experiment. Robust standard errors clustered on the team level in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

Table A9: Heterogeneous Treatment Effects by Employee Skill Level

	Num. of Trainings	
	(1)	(2)
Treatment	−0.194**	−0.171*
	(0.091)	(0.092)
Treatment × High-Skill		−0.101
		(0.066)
High-Skill		0.115**
		(0.053)
Constant	0.374***	0.348***
	(0.077)	(0.077)
p-value High-Skill		0.062
Observations	2582	2234

Note: Table shows estimates from linear regressions of the number of trainings on a treatment indicator interacted with the 'High-Skill' dummy. The variable 'High-Skill' takes value 1 for all employees with an above median average assessment score prior to the experiment and value 0 otherwise. Robust standard errors clustered on the team level in parentheses. *, **, *** indicate significance on the 10%, 5% and 1% level, respectively.

A.12 Classification of Work Assignments

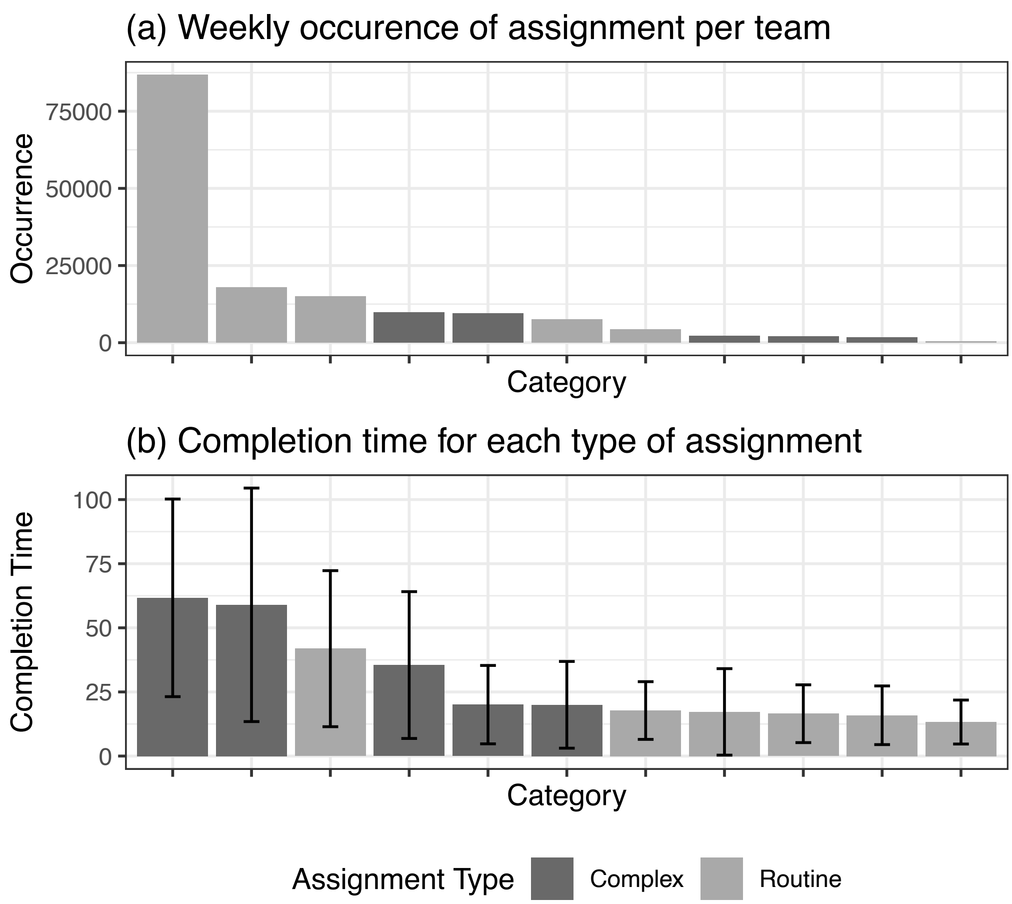

Figure A6: Classification of Work Assignments by Complexity

A.13 Treatment Effects by Training Purpose and Employee Type

Table A10: Treatment effects by Training Purpose and Employee Type

	Skill Preservation		New Skill Acquisition	
	High-Skill	Low-Skill	High-Skill	Low-Skill
Treatment	−0.095	−0.108*	−0.185**	−0.070
	(0.076)	(0.058)	(0.071)	(0.050)
Constant	0.226***	0.196***	0.235***	0.153***
	(0.057)	(0.051)	(0.066)	(0.040)
Observations	1043	1191	1043	1191

Note: Table shows estimates from linear regressions on trainings devoted to skill acquisition (columns 1 and 2) and skill preservation (columns 3 and 4). The variable 'High-Skill' takes value 1 for all employees with an above median average assessment score prior to the experiment and value 0 otherwise. Robust standard errors clustered on the team level in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

A.14 Post-Experimental Survey Items (Translated)

Item Text	Name
How useful are the following elements of the learning platform to you? • The skill assessments on the platform • The development ideas you can write down after the skill assessment	*Usefulness*
How easy was the handling of the following elements of the platform? • The skill assessments on the platform • The development ideas you can write down after the skill assessment	*Handling*
I participate in the skill assessment because…	
… it is my job.	*Obligation*
… my manager is expecting it from me.	*Demand*
… it helps me to find out how I can improve.	*Improvement*
… to demonstrate to others that I am compentent.	*Competence*
If I do not participate in the skill assessment, it is because…	
… I do not have the time to participate.	*Time*
… the skill is not relevant for my daily work.	*Mismatch*
… I already know that I will achieve a high score.	*Overconfidence*
… I expect to perfom poorly.	*Incompetence*
Thinking back to the last 6 months: How often did you have a meeting with your manager to talk about your skill development?	*Number of Meetings*

Item Text	Name
Thinking back to the last 6 months, how do you evaluate the following statements regarding your development process? • My manager helps me find the right training measures. • My manager supports me in learning independently. • My manager is interested in what I am currently learning. • My manager helps me try out new solutions, even if I make mistakes in doing so. • Every time I have a good idea for my own development, my managers implements it. • With regards to my development, it often happened that I was presented with a fait accompli by my manager. • I always had a say in the decisions about my own development.	*Supervisor Support*
How satisfied are you with your work at the moment?	*Satisfaction*
I often think about changing my job.	*Turnover*
• My supervisor gives me special recognition when my work performance is especially good. • My supervisor always gives me positive feedback when I perform well.	*Positive Feedback*
• My supervisor shows me his/her displeasure when my work is below acceptable standards. • My supervisor lets me know about it when I perform poorly.	*Negative Feedback*

<table>
<thead>
<tr><th>Item Text</th><th>Name</th></tr>
</thead>
<tbody>
<tr><td>Now we ask you to assess what leadership skills you think your manager has. We have included a short description of a person who has this competence.
Leadership and coaching: adapts leadership style to the current environment, motivates and mentors others.
Influence on others: is well connected, uses persuasion and their own authority to achieve team goals.
Interpersonal relationships: responds appropriately to the needs of others, understands how to give good feedback, builds trust with others quickly.
Conflict management: anticipates conflicts and complaints and resolve them constructively.</td><td>Leadership Skills</td></tr>
<tr><td>I value the professional skills of my manager.
I respect my manager’s knowledge and expertise.
I am impressed by my manager’s knowledge of the job.</td><td>Perceived Competence</td></tr>
</tbody>
</table>