The Janus Face of Artificial Intelligence Feedback: Deployment Versus Disclosure Effects on Employee Performance

Siliang Tong
Nanyang Business School, Nanyang Technological University

Nan Jia
Marshall School of Business, University of Southern California

Xueming Luo
Temple University

Zheng Fang
Business School, Sichuan University

Nanyang Business School Research Paper No. 21-40

Nanyang Business School
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798
https://www.ntu.edu.sg/

The Janus Face of Artificial Intelligence Feedback: Deployment Versus Disclosure Effects on Employee Performance

Siliang Tong (Nanyang Technological University)

Nan Jia (University of Southern California)

Xueming Luo (Temple University)

Zheng Fang (Sichuan University)

Abstract

Companies are increasingly using artificial intelligence (AI) to provide performance feedback to employees, by tracking employee behavior at work, automating performance evaluations, and recommending job improvements. However, this application of AI has provoked much debate. On the one hand, powerful AI data analytics increase the quality of feedback, which may enhance employee productivity ("deployment effect"). On the other hand, employees may develop a negative perception of AI feedback once it is disclosed to them, thus harming their productivity ("disclosure effect"). We examine these two effects theoretically and test them empirically using data from a field experiment. We find strong evidence that both effects coexist, and that the adverse disclosure effect is mitigated by employees' tenure in the firm. These findings offer pivotal implications for management theory, practice, and public policies.

Managerial abstract

Artificial Intelligence (AI) technologies are bound to transform how companies manage employees. We examine the use of AI to generate performance feedback for employees. We demonstrate that AI significantly increases the accuracy and consistency of the analyses of information collected, and the relevance of feedback to each employee. These advantages of AI help employees achieve greater job performance at scale, and thus create value for companies. However, our study also alerts companies to the negative effect of disclosing using AI to employee that results from employees' negative perceptions about the deployment of AI, which offsets the business value created by AI. To alleviate value-destroying disclosure effect, we suggest that companies be more proactive in communicating with their employees about the objectives, benefits, and scope of AI applications in order to assuage their concerns. Moreover, the result of the allayed negative AI disclosure effect among employees with a longer tenure in the company suggests that companies may consider deploying AI in a tiered instead of a uniform fashion, i.e., using AI to provide performance feedback to veteran employees but using human managers to provide performance feedback to novices.

Keywords: Artificial Intelligence, Employee Productivity, Employee Performance Feedback, Employee Perceptual Bias, Field Experiment, New Technology in Management

Acknowledgement: The authors would like to thank the anonymous company for sponsoring the field experiment, as well as the participants of the Organizational Behavior Lab of the USC Marshall School, strategy seminar of the University of Colorado Boulder, the Conference on Artificial Intelligence, Machine Learning, and Business Analytics, the AI and Strategy Virtual Workshop, the MIT CODE conference, the Strategy Science Conference, research workshop of the Chinese University of Hong Kong-Shenzhen, research seminar of the Temple University, for helpful comments. All errors are our own. The corresponding author is Zheng Fang.

Siliang Tong is Assistant Professor of Information Technology and Operation Management, Nanyang Business School, Nanyang Technological University, Singapore (email: jack.tong@ntu.edu.sg)

Nan Jia is Associate Professor Management and Organization, Marshall School of Business, University of Southern California, USA (email: nan.jia@marshall.usc.edu)

Xueming Luo is Charles Gilliland Distinguished Professor of Marketing, MIS, and Strategy, and Director and Founder of Global Institute for Artificial Intelligence and Business Analytics, Fox School of Business, Temple University, USA (email: luoxm@temple.edu)

Zheng Fang is Professor of Marketing, Business School of Sichuan University, Sichuan University, China (email: 149281891@qq.com)

INTRODUCTION

Artificial Intelligence (AI) is playing an increasingly important role in firm management (Agrawal, Gans, & Goldfarb, 2018; Lee, 2018;Luo et al. 2021; Iansiti and Lakhani 2020). A burgeoning area of AI application is to conduct job evaluation and provide performance feedback to employees. Leveraging big data analytics and self-learning capabilities, AI applications can track employees' activities at work, evaluate job performance, and generate recommendations for changes that could improve employee productivity. For example, Enaible, an entrepreneurial platform, has developed an AI program that tracks employees' work remotely. The AI program assesses each employee's typical workflow, assigns a "productivity score," and identifies ways to increase efficiency in the workflow. This AI feedback program has been licensed to the Dubai customs agency and Omnicom Media Group, and is allegedly in late-stage talks with Delta Airlines and CVS Health (Heaven, 2020). MetLife, a leading insurance company, uses an AI training program to track service employees' conversations with customers and make recommendations to employees on what to improve on the job (Roose, 2019). Unilever adopts AI programs to provide feedback to new employees and help them better settle into the job (Marr, 2018).

Using AI to provide performance feedback in the workplace has provoked much debate. On the one hand, advanced data analytics enable AI to comprehensively track employees' behavior on the job, accurately assess their productivity, and generate personalized recommendations for job improvement, all in a consistent and accurate manner (Heaven, 2020). These features are thought to help employees improve their job performance at scale (Colangelo, 2020). On the other hand, there exists a concern that implementing AI programs, especially without a transparent policy, might tilt the power balance against employees (Bughin & Manyika, 2019). Employees may develop a negative perception about AI as a management tool once it is disclosed because workplace surveillance can undermine trust and damage morale (Cheatham et al., 2019; Premuzic et al., 2018), thus hindering employee performance (Carpenter, 2019; Roe, 2018; Teich, 2019). Therefore, companies face a trade-off when adopting AI to generate performance feedback. If they disclose the AI feedback to their employees, they will not reap its full value as a management tool. However, employees have the right to know they are being monitored by machines and algorithms, requiring regulations that mandate the disclosure of AI usage in firms (MacCarthy, 2020; O'keefe et al., 2019). Nevertheless, beyond anecdotes and industrial reports, this trade-off has not been examined sufficiently in the academic literature in a systematic manner. As a result, there is a lack of systematic understanding about whether and why AI feedback improves or harms employee performance, and how companies can reduce the potential negative effects of disclosing the use of AI to their employees.

We aim to address this knowledge gap. First, drawing on prior research on the general abilities of AI in data mining and analytics (Davenport & Ronanki, 2018; Luo et al., 2019; Thiel, 2019),[1] we argue that, relative to human managers, AI feedback can consistently analyze a larger amount of data with greater precision, which increases the accuracy of the evaluations of employee performance. AI feedback is also more relevant to individual employees because AI can achieve a higher level of customization. Both factors contribute to higher quality feedback which, in turn, leads to greater employee productivity; we refer to this as a positive "deployment effect." Second, we draw on the research on negative human perceptions against AI (Huang & Rust, 2018; Longoni, Bonezzi, & Morewedge, 2019; Newman, Fast, & Harmon, 2020; for a review, see Glikson and Woolley, 2020) and research on AI's effect on the displacement of labor (Acemoglu & Restrepo, 2020; Agrawal, Gans, & Goldfarb, 2019), to argue that employees lack trust in AI feedback and are concerned about the risk of replacement by AI. Both concerns will reduce employees' productivity once they are informed of the act of using AI—relative to human managers—to generate feedback to them; we refer to this as a negative "disclosure effect."

We exploit data from a field experiment conducted by a large financial services company, in which 265 employees were randomly assigned to receive performance feedback generated by AI or

[1] Because AI technologies nowadays can perform well-defined, structured tasks (e.g., Brynjolfsson and Mitchell, 2017; Luo et al. 2019), we examine the use of AI to provide feedback for jobs with structured tasks and provide more detailed discussion of this scope condition subsequently.

human managers. A novel feature of the experimental design is that the disclosure of feedback providers' identities is also randomized. That is, a human manager could be disclosed either as such or as the AI, and the AI could be disclosed as such or as a human manager. This simple experiment design randomizes AI "disclosure" independently of the randomization of "deployment," as a result of which it effectively isolates the deployment effect from the disclosure effect of AI feedback on employee performance.

Overall, we find a positive net effect of providing AI feedback on employee performance than providing human feedback. However, the "deployment effect" and "disclosure effect" are opposite in direction once they are teased apart from each other. The results support a positive deployment effect of AI feedback; employees who received AI feedback attained 12.9% higher job performance than those receiving human managers' feedback. That is, AI feedback improves employee performance more than human feedback does. Further, we find that AI generates higher quality feedback for employees than human managers do in that AI identifies more mistakes and makes more recommendations to correct each mistake in order to improve employees' job skills. The results of our mediation analysis provide suggestive evidence that AI deployment provides higher quality feedback, which then increases employees' learning and job performance.

In contrast, we find a negative disclosure effect; employees informed of receiving feedback from AI achieved 5.4% lower job performance than those informed of receiving feedback from human managers. The survey results show that disclosing AI feedback induces employees to develop negative perceptions about AI feedback, including lower trust in the quality of the feedback and higher concerns over job displacement risk. Our mediation analyses offer suggestive evidence that disclosing AI feedback reduces employees' trust in the quality of feedback and heightens their job displacement risk, both of which harm their learning and job performance.

Moreover, to better understand how to mitigate the adverse disclosure effect, we examine how this is affected by employee heterogeneity. We find that the negative disclosure effect is attenuated by employees' tenure in the firm. This is consistent with our theory that longer-serving employees often have extensive networks and relational capital within the firm, providing resources and support against adverse shocks. As a result, the negative disclosure effect of AI feedback is less severe for employees with a longer tenure than it is for those with a shorter tenure.

This study contributes to emerging research on how AI technologies help shape business management. To the best of our knowledge, this study is among the first to examine the new and important phenomenon of using AI to generate employee performance feedback in the workplace. Advances in deep learning and neural network techniques empower AI to perform the managerial task of feedback provision, which entails not only tracking employee performance but also generating customized performance evaluations and personalized recommendations to improve employees' job skills at scale. This represents an unprecedented opportunity for firms to create value (Bughin & Manyika, 2019). Thus, this paper takes an initial step in extending prior research on AI applications in production and marketing (Aghion et al., 2017; Aron et al., 2011; Brynjolfsson et al., 2019; Schanke et al., 2020; Sun et al., 2019) to investigate the role of AI in managing employees, particularly at the interface between AI and employees.

Second, we address the benefits and costs trade-offs inherent in using AI to provide feedback by developing a theoretical basis for a value-enhancing "deployment effect" and value-destroying "disclosure effect." While deploying AI to generate feedback creates value, we show that, paradoxically, disclosing AI feedback—regardless of the true feedback identity—reduces employee performance. We theoretically and empirically unravel the coexistence of the two effects, hence revealing the "Janus face" of AI feedback. It is crucial to distinguish between these countervailing effects because, without accounting for the disclosure effect created by employees' subjective perceptions against AI, research may substantially underestimate the true value of AI and overlook opportunities to mitigate factors that hamper its potential to create firm value. In this sense, we advance the literature by quantifying the degree to which the AI disclosure effect can offset the productivity gain of AI deployment.

Furthermore, our findings offer valuable implications for firms. Global investment in AI reached USD 35 billion in U.S. dollars in 2019 and is projected to double within the next two years (Deloitte, 2019). Our findings suggest that AI feedback improves employee performance beyond that of human feedback, suggesting substantial business returns from investing in AI applications in firm management. Despite the negative disclosure effect of AI feedback, its net effect on employee performance is positive; the magnitude of the value-enhancing deployment effect exceeds that of the value-reducing disclosure effect. However, our results on the negative disclosure effect indicate that firms need to be aware of employees' negative perceptions. Here, we recommend several strategies that companies can use to alleviate these negative effects. In particular, our finding that the negative AI disclosure effect decreases with employee tenure means companies may consider using AI technologies to provide feedback to veteran employees and using human managers to provide feedback to novice employees. This combination may allow firms to reap even higher returns on their AI investment.

Finally, our study generates critical public policy implications. The proliferation of AI in the workplace has attracted the attention of policymakers who are concerned that AI may jeopardize employee wellbeing, resulting in regulations that increase the transparency of AI usage (Martinez, 2019). Our findings show that disclosing AI feedback increases negative perceptions among employees, thus reducing employee productivity. Therefore, disclosing AI deployment must be accompanied by measures that address workers' negative perceptions of AI. These measures may include providing information on how AI functions and increasing societal support by means of subsidies and employee retraining. Therefore, it requires a portfolio of policies that tackle a range of related issues, instead of an unidimensional policy on the transparency of AI alone, in order to enable AI technologies to provide more benefits for both firms and their employees.

THEORY AND HYPOTHESES

Conceptual Background of Employee Performance Evaluation and Feedback

For over a century, it has been known in management theory that accurate information on how much and how well employees work constitutes a critical path to higher productivity and firm value (Taylor, 1911). In this sense, employees' performance evaluation and feedback, which entails collecting information about their behavior on the job, assessing their job performance, and providing feedback to employees on what needs to be changed in order to improve their performance (Latham & Kinne, 1974; Oldham & Cummings, 1996), is a crucial part of firm management. These activities lie at the heart of the "information role" of managers, which requires that managers monitor the workplace, including employees, to generate, process, and disseminate information to members of the firm (Mintzberg, 1990). In data analytics, AI technologies are used to make accurate and comprehensive predictions (Agrawal, Gans, & Goldfarb, 2016; Huang & Rust, 2018), suggesting that AI has the potential to perform these information functions. Indeed, firms constantly adopt technologies to automate the process of labor-intensive mechanical jobs. For example, Amazon uses algorithms to evaluate the performance of its warehouse employees (Ip, 2019). However, it has become increasingly popular for firms to use cutting-edge AI technologies to evaluate and provide feedback to workers as well, as occurring in the aforementioned examples of Enaible, MetLife, and Unilever.

Technical Advantages: Productivity-Enhancing AI Deployment Effect

In performing structured and well-defined tasks, AI technologies have superior data analytics skills compared with those of humans, thus enabling AI to make more accurate predictions (Jarrahi, 2018; Verma & Agrawal, 2016). These advantages have been shown by prior research to improve firm value. Specifically, some studies focus on how AI assists firms in serving external stakeholders, particularly customers, by producing higher quality products and services and reducing costs. For example, using AI in medical diagnoses reduces errors (Aron et al., 2011; Meyer et al., 2014), AI-powered chatbots increase customer purchases (Luo et al., 2019), AI-based translation software delivers faster and cheaper translation services (Brynjolfsson et al., 2019), and AI applications in R&D boost the drug discovery process (Fleming, 2018). Others have examined how AI can be used internally in firm management to create value; for example, Bai et al (2020) show that AI can assign tasks to warehouse employees to increase work efficiency.

Drawing on the abovementioned advantages of AI technologies, we argue that deploying AI (versus human managers) to provide performance feedback to employees on jobs with well-structured tasks increases their performance for two reasons. First, AI is able to rapidly analyze a large amount of data on employees' activities and behavior with greater precision, thereby increasing the accuracy of performance assessments. As noted earlier, accurate information on how much and how well employees work has traditionally been valued in firm management as a critical path to higher job productivity (Taylor, 1911). In contemporary businesses, data analytics have both grown in prominence and become more challenging because of the proliferation of data available for analyses. Rapid advancements in hardware and software enable firms to capture a larger amount of data and a great variety including unstructured data such as text, audio, and video (Verma & Agrawal, 2016). When analyzing large and complicated data, algorithms and computer programs generate results that are more accurate than those of humans (Jordan & Mitchell, 2015; Tarafdar et al., 2019; Whitt, 2006). Moreover, AI can assess large data more comprehensively than humans can, because AI can draw on a much larger training data set which contains both successful and failed precedents, than that available in human memory. In other words, AI can increase the quality of employee performance assessments than human managers can, by more accurately and more speedily analyzing a wider range of data on how employees perform on the job.

Second, compared with human managers, AI can generate recommendations that are more relevant for each employee on jobs with well-structured tasks. The ability of AI analytics to analyze enormous quantities of data deeply and quickly enables it to generate "personalized" recommendations at scale, that is, to make accurate and individualized recommendations (Agrawal, 2018; Huang and Rust, 2018). While human managers can also make personalized recommendations, their cognitive limits constrain the speed at which they process data as well as their ability to achieve this goal for a large number of cases. In other words, AI increases the relevance of the feedback provided to each employee by more accurately addressing each employee's unique situation and challenges on the job. In contrast, human managers have limited attention and capacity, which hinders them from providing highly "customized" feedback for a large number of employees in a consistent and accurate fashion (Brynjolfsson and Mitchell, 2017; Luo et al., 2019).

Based on these two technical advantages of AI, we propose that the deployment of AI in providing performance feedback to employees on jobs with well-structured tasks generates a positive effect on employees' job performance compared with deploying human managers to provide such feedback, which we refer to as a positive "deployment effect."

H1 (deployment effect): For jobs with well-structured tasks, deploying AI instead of human managers to provide performance feedback to employee has a positive effect on employees' job performance.

In developing H1, we have discussed that AI feedback is of higher quality in that its assessments more accurately capture employees' performance on jobs with well-structured tasks, and its recommendations are more relevant to the unique situation of each employee. The content of the feedback conveyed critically shapes how much the feedback is accepted by the recipient (for a review, see Wisniewski, Zierer, & Hattie, 2020). The main goal of performance feedback is to eliminate the discrepancy between the subject's current understanding and the performance goal (Sadler, 1989). As such, providing feedback that more accurately captures this discrepancy makes it more likely that the feedback will be accepted and learned by the employees. In other words, employees learn more from the recommendations generated by AI than from those of human managers, because the former are of higher quality and more relevant (Latham & Kinne, 1974). This learning then improves performance, since the recommendations more accurately capture the discrepancy between an employee's current behavior and what she needs to do to attain higher performance (Oldham & Cummings, 1996). Thus, AI feedback is of higher quality than human feedback, which, in turn, improves employees' learning and job performance. Therefore, we propose the following underlying mechanism for the deployment effect:

H2: The positive deployment effect of AI feedback on employee performance as captured by H1 is mediated by higher quality of AI feedback than human feedback, which in turn results in a higher level of employee learning from AI feedback than from human feedback.

Negative Perceptions: Productivity-Destroying AI Disclosure Effect

Thus far, our theory focuses on the technical advantages of AI in providing performance feedback on jobs with well-structured tasks to employees. However, once such use of AI in this regard is disclosed to employees, they may experience “algorithm aversion.” This concept refers to people holding negative perceptions about the recommendations and decisions made by algorithms, regardless of the content quality, relative to those made by other people (e.g., Kahneman, 2011). For example, Dietvorst, Simmons, and Massey (2015) show that users are less tolerant of forecast errors made by algorithms than those made by humans. Patients are less receptive to AI medical assistance, citing a lack of uniqueness (Longoni et al., 2019b). Furthermore, customers are less welcoming to AI chatbots (Luo et al., 2019) and humanoid robots (Leung et al., 2018; Mende et al., 2019). In general, workers may not trust AI algorithms, despite their technical advantages (for a review, see Glikson & Woolley, 2020).

Prior research demonstrates that employees develop negative perceptions about being managed by AI because they regard the tracking and surveillance by AI at work as an infringement of their privacy (Raveendhran & Fast, 2019), construe using AI in management as lacking procedural justice (Newman et al., 2020) and consider AI as undermining their sense of autonomy at work (Möhlmann & Zalmanson, 2017a). Thus, based on this stream of research, we argue that these negative perceptions likely reduce employees’ trust in the disclosed AI feedback, which adversely affects their learning from the feedback and subsequent job performance.

Moreover, there exists a body of literature on the risks of job displacement by AI technologies in labor markets(Acemoglu & Restrepo, 2018, 2020; Ajay Agrawal et al., 2016; Webb et al., 2019). At the individual level, this means that employees are concerned about or even fear being replaced by AI technologies (Garimella, 2018; Felten, Raj, & Seamans, 2019). Although using AI to generate employee performance feedback does not replace employees (rather, AI feedback may replace human managers' feedback), the general fear of AI's job displacement effect among employees may generate negative spillover onto how they perceive AI feedback (Roose, 2019). Moreover, employees may worry about the firms' moral hazard, suspecting that the information collected by AI about their job behavior may be used against them later, perhaps to sabotage them or to replace them, which may be demoralizing (Roe, 2018; Makridakis, 2017; Agrawal, Gans, et al., 2019).[2]

Overall, employees' negative perceptions (i.e., lower trust in the quality of feedback and higher concerns over job replacement risk) of disclosed AI feedback likely harm their learning from the feedback, thus reducing their performance, compared with disclosed human feedback. Therefore, we propose that, all else being equal, disclosing AI feedback (versus human managers' feedback) to employees decreases their performance, which we call the negative "disclosure effect."

H3 (disclosure effect): For jobs with well-structured tasks, disclosing to employees that performance feedback is provided by AI instead of human managers has a negative impact on employees' job performance.

In developing H3, we have posited that the disclosure of AI feedback reduces employees' performance through the following two mechanisms. First, employees' lower trust in the disclosed AI feedback means they are less likely to accept its feedback and follow its recommendations, resulting in lower employee learning and performance (e.g., Wisniewski et al., 2020). Second, the fear of being replaced by AI demoralizes the employees (e.g., Ashforth, 1994), which also reduces their motivation to learn from AI feedback and thus harms their performance. Therefore, the disclosure of AI feedback first induces lower trust in the quality of the feedback and higher perceived job displacement risk, both of

[2] Recent research has started to examine the possibility of "algorithm appreciation," which refers to people's greater adherence to the advice given by algorithms than that by other persons. Logg, Minson and Moore (2019) argue that individuals "may feel more comfortable with algorithmic advice in domains that feature a concrete, external standard of accuracy, such as investment decisions or sports predictions" (Page 91), or in context where algorithms have been historically applied such as weather forecasts. Our context features neither situation.

which in turn reduce employees' learning and job performance. Hence, we propose the following as the underlying mechanisms for the disclosure effect.

H4: The negative disclosure effect of AI feedback on employee performance as captured by H3 is mediated by employees' lower trust in the quality of AI feedback and higher perceived job displacement risk by AI, both of which in turn result in a lower level of employee learning from AI feedback than from human feedback.

Employee Tenure Alleviates the Negative Disclosure Effect

Because we aim to understand how to increase the value of AI technologies in firm management, we investigate circumstances that may alleviate the negative consequences of the AI disclosure effect. We argue that this negative effect is less severe among employees with a longer tenure in the firm than it is for those with a shorter tenure. Specifically, employees who have worked longer in a firm often have stronger and more extensive networks within the firm, or relational capital (Hunt & Saul, 1975; Perry & Mankin, 2004). Greater relational capital provides resources and support through mechanisms such as reciprocity, which safeguard employees from adverse shocks (Rogan & Mors, 2014). Indeed, employees with a longer tenure may perceive that they are better "protected" in the firm (Ewert, 1984; Webster, 1993), thus likely alleviating them from the adverse effects of disclosing AI feedback. In other words, relative to those with a shorter tenure, employees with a longer tenure are likely to develop perceptions that are less negative in response to the disclosure of AI feedback, because they have accumulated more extensive networks and stronger relational capital within the firm. As a result, we posit that the negative effect of AI disclosure on performance decreases in severity with an increase in employee tenure.

H5: For jobs with well-structured tasks, the negative disclosure effect of AI feedback on employee performance is less pronounced among employees who hold a longer tenure in the firm.

FIELD EXPERIMENT SETTING AND DESIGN

Company Setting

Omega Corp (we use a pseudo name to ensure anonymity) was a large financial services company in Asia, with over 12 million customers. The company offered a broad set of financial products, including personal lending, bridge loans, refinancing, and equity investment. Because the personal loan business grew significantly in the local market, the company had a large call center that promoted its financial products and collected overdue payments from delinquent borrowers. On average, individuals borrowed $2,000 on a 12-month installment, mainly to purchase products such as mobile phones, TVs, computers, and household furniture. The booming personal loan business engendered a substantially high rate of overdue and default payments. Therefore, many employees were hired in the call center to collect payments that were overdue by more than 15 days. Furthermore, to improve the employees' job performance in collecting such payments, the company conventionally relied on experienced human managers in the quality control department to provide feedback on their calls. Specifically, human managers evaluated employees' collection calls to find mistakes that should be rectified and provided recommendations to employees to improve their job skills in collecting overdue payments from delinquent borrowers. All human managers had extensive loan collection experience and feedback provision skills, according to the company.

With the advent of new AI technologies, Omega Corp worked with a leading technology platform to deploy an AI system to evaluate the collection calls and provide employees with job feedback. This AI system was enabled by state-of-the-art deep learning neural network-based speech analytic algorithms and trained with an enormous, archived data set of recorded collection calls and human managers' feedback recommendations from similar firms in the industry. Essentially, the AI feedback system comprised four key components. First, its automatic speech recognition (ASR) component converted phone call conversations between employees and customers from unstructured audio data into text scripts. Second, the natural language understanding (NLU) component conducted semantic parsing to embed the scripts into numerical representations. Third, its hypothesis searching (HS) component applied machine learning models (i.e., Word2Vec) to calculate the distance score between the best-practices in the knowledge bank and the scripts of the employee to determine the employee's effectiveness in persuading customers. That is, it analyzed the calls to find mistakes that should be rectified for each employee; it automated job performance evaluations. Fourth, the feedback recommendation (FR) component generated comprehensive and personalized recommendations to remedy each mistake made by the employee (i.e., to improve her job performance). In other words, because the AI system was powered by deep learning speech recognition, speech-to-text, and semantic parsing technologies, it was able to automate the overall feedback process when monitoring employee job performance. Note that this AI feedback was highly advanced because it captured comprehensive information of employees' calls with customers, based on which evaluated these calls and providing recommendations to improve productivity. In this sense, it could function as a management tool to free human managers from the routine, repetitive tasks of assessing subordinates' calls, identifying their mistakes, and making suggestions for correction.

This AI feedback system was highly attractive to the company because it provided accurate information on how much and how well each employee worked and effective recommendations to improve their productivity. The pilot testing period showed that the system was competent, with a very low error rate (less than 1%) when identifying employees' mistakes. Omega Corp had a keen interest in designing the experiment to quantify the effect of deploying and disclosing AI feedback on its employees' job performance. The research team gained access to the field experiment data and conducted the analyses together with the company.

Experimental Design

The field experiment followed a two-by-two full-factorial design, with the two dimensions being the deployment of AI or human feedback and the disclosure of AI or human feedback. A novel feature of this experiment design is that the randomization of AI feedback disclosure is independent from the randomization of AI feedback deployment. Specifically, the company randomly selected 265 full-time employees who joined the firm recently and were still in the probation period (to minimize within-group performance variations caused by diverse working experience) to four experiment groups as shown in

Figure 1, to receive performance feedback from the AI system or human managers *and* to be informed that they receive AI or human feedback. The first condition (Group 1, N = 64) was "Feedback Generated by and Feedback Provider Disclosed as AI," in which employees received feedback provided by the AI system and were informed as such. In the second condition (Group 2, N = 69), called the "Feedback Generated by AI but Feedback Provider Disclosed as Human Managers," wherein the AI system generated feedback for each employee but employees were informed that human managers generated the feedback. In the third condition (Group 3, N = 66), called the "Feedback Generated by Human Managers but Feedback Provider Disclosed as AI," human managers generated the feedback but employees were informed that the feedback was provided by the AI system. In the fourth condition (Group 4, N = 66), "Feedback Generated by and Feedback Provider Disclosed as Human Managers," employees received feedback from human managers and were informed as such.

This experiment design effectively separates the effect of deployment from that of the disclosure of AI feedback. Specifically, as illustrated in Figure 1, holding constant the disclosed identity of the feedback provider as AI, we can gauge the *deployment* effect of AI feedback as Δ_{13} (the performance difference between Group 1 and Group 3). Similarly, holding constant the disclosed identity of the feedback provider as human managers, we can measure the deployment effect of AI feedback as Δ_{24} (the performance difference between Group 2 and Group 4). We can also analyze the *average* deployment effects of these two differences by comparing the pooled Groups 1 and 2 (in both groups, the feedback was generated by AI) with the pooled Groups 3 and 4 (in both groups, the feedback was generated by human managers).

Furthermore, holding constant the deployed feedback provider as AI, we can measure the *disclosure* effect of AI feedback as Δ_{12} (the performance difference between Group 1 and Group 2). Similarly, holding constant the deployed feedback provider as human managers, we can gauge the disclosure effect of AI feedback as Δ_{34} (the performance difference between Group 3 and Group 4). We also analyze the average disclosure effects of these two differences by comparing the pooled Groups 1 and 3 (both are informed that AI provides feedback) with the pooled Groups 2 and 4 (both are informed that human managers provide feedback).

In a traditional randomized control trial with AI deployment as the treatment, employees may also know the status of this treatment, which causes the disclosure effect to confound the deployment effect. In contrast, in our data the manipulation of treatment parcels out the disclosure effects from the deployment effects. Indeed, it is pivotal to disentangle these two effects in order to accurately measure the true value of AI feedback. Companies can use this experimental method to gauge the true effects of deploying and disclosing AI feedback in order to properly budget for the AI investment.

During the month-long experiment, each employee was required to work the same load (completing 100 collection calls per day) to rule out the alternative explanation of differing workloads. Moreover, all employees assigned to the four experiment conditions had the same work schedules, that is, the distributions of the workloads on workdays and weekends, and between mornings and afternoons on a working day were the same for these employees. Furthermore, all employees received a randomly assigned list of delinquent customers to call each day from the call center system and they needed to call each customer in the list in the sequence that was specified in the list, which helps rule out the alternative explanations that are created by concerns over customer heterogeneity. A total of six seasoned managers were selected randomly from the quality control department as human feedback providers in the experiment.[3] On average, each manager monitored 22 employees during the experiment period and randomly selected five phone calls per employee each day to evaluate and provide recommendations. Thus, each manager monitored about 110 random calls per day, which was equivalent to their normal workload in the company. No employee or human manager who participated in the experiment left the company during the experiment; hence, there exists no survival bias in our data.

[3] Omega Corp allocated to the managers who were involved in the experiment only the function of training employees or providing data-driven performance feedback to employees, whereas the Human Resource department retained the decisions over employees' promotion/demotion/termination. In other words, these managers were relieved of a number of standard managerial functions (such as promoting or terminating employees), and their fundamental goal was to solely provide effective feedback to train workers.

The AI system and the human managers in our experiment performed the same feedback task, which was to listen to the collection calls, identify mistakes, and make personalized recommendations retrieved from the company knowledge bank to rectify each mistake (see the online Appendix A for some examples). Across the four experiment conditions, employees received daily feedback emails sent by the company's quality control department. To rule out alternative explanations, the AI and human feedback was provided in the same format. Each started with the disclosed identity of the feedback provider, either as "a manager in the quality control department" (without specifying which manager did that; this approach rules out the potentially confounding effects of managerial heterogeneity such as their popularity among employees) or as the "the AI feedback system in the quality control department." This information was followed by reproduced scripts of the calls made by the employee the previous day that contained mistakes, with the text of each mistake highlighted and followed by recommended alternative scripts to correct it. For example, mistakes included using inappropriate persuasion strategies, providing incomplete or vague information, or insulting the customer with aggressive and emotional expressions.

The company ensured that the employees and human managers in the experiment had no knowledge about the true identity of the feedback provider (other than what was disclosed to employees), guaranteeing that nobody could strategically respond to the treatment manipulation. For example, a strategic manager could start to provide lower-quality feedback if she knew it would be disclosed to employees as coming from the AI system. Further, to maintain the confidentiality of the information on job performance, company policy forbade employees and managers from sharing the feedback and job performance with coworkers (this policy persisted before the experiment). This practice reduces the possible concerns related to spillover effects and contamination across experiment conditions.

Note that in this experiment, the company restricted the AI system to assessing the same number of calls (five calls per employee per day) as human managers, although the AI system can analyze far more calls per employee than human managers can. This restriction makes our results more conservative.[4] None of the employees in the experiment had prior experience of receiving feedback from any AI system. We also control for the identity of the human manager who provided feedback to each employee in the month prior to the experiment in the following data analyses.

Data and Randomization Check

As shown in Panel A of Table 1, on average, the full-time employees in our experiment are 20 years old, 42% have received post high school education, and 7.5% had worked in a call center prior to joining Omega Corp. Their average tenure at Omega Corp is three months, consistent with the fact that these are recently hired employees who are still in the probation period and thus need a substantial amount of training feedback to improve their job performance. In the month prior to the experiment, their average collection amount was 9,540 in local currency (USD 1,360). We conduct a randomization check of these variables and report the results in Panel B of Table 2. A one-way analysis of the variance (ANOVA) and chi-square test fail to reject the null hypothesis that the mean values of these variables are not different among the four experiment conditions. Thus, the data pass the randomization check. Panel C reports the summary statistics of the employees' job performance and other key variables.

MODELS AND RESULTS

Model-Free Evidence

Figure 2 presents the unconditional mean value of employees' job performance for the four experiment groups. Employee's job performance is measured as the average payment collected by each

[4] While AI indeed has greater capacity to train more employees within the same time frame than human managers, we constrain this advantage by letting AI assess the same number of calls and provide feedback to the same number of employees as human managers are accustomed to at work, because we consider it important to distinguish the performance of AI as providing higher quality training given the same amount of information (or higher efficiency), from the possibility of providing lower quality training but more cheaply on a larger scale. Our theory development focuses on the reasons why AI can provide high quality training.

employee during the experiment month.[5] First, we examine the deployment effect of AI feedback. Specifically, we compare the performance of the employees in Groups 1 and 3, all of whom are told that the AI system evaluates their job performance and provides feedback to them. However, while Group 1 indeed receives feedback generated by the AI, Group 3 actually receives feedback generated by human managers. Thus, the performance difference between the two groups captures the AI deployment effect, holding constant the disclosed identity of the feedback provider as the AI. The average job performance of the employees in Group 1 (i.e., daily collection amount 11,211.891 in local currency) is 12.2% higher than that of the employees in Group 3 (i.e., daily collection amount 9,994.130 in local currency). In the same vein, we compare the performance of the employees in Groups 2 and 4, who are all informed that human managers evaluate their job performance and provide feedback to them. However, Group 2 actually receives the feedback generated by the AI, whereas Group 4 receives feedback generated by human managers. Therefore, their difference captures the AI deployment effect, holding constant the disclosed identity of the feedback provider as a human manager. The average job performance of the employees in Group 2 (i.e., daily collection amount 11,881.887 in local currency) is 13.3% higher than that of employees in Group 4 (i.e., daily collection amount 10,488.515 in local currency). Therefore, regardless of the disclosed feedback identity, employees who receive feedback that is actually generated by the AI always outperform those receive feedback that is actually generated by human managers. As shown in online Appendix B, the model-based regression results of the treatment groups and a host of controlling covariates are highly consistent with the model-free evidence. These results provide preliminary evidence for H1 regarding the positive deployment effect of AI feedback on employee performance.

Next, we examine the disclosure effect of AI feedback. We first compare the performance of those in Groups 1 and 2, who all receive the feedback generated by AI but are informed differently: the

[5] Here, we focus on monthly performance in the analyses because in practice Omega Corp evaluates employee performance on a monthly basis. In additional analyses, we explore the dynamic effect by splitting the data into the first 15 days and the second 15 days of the experiment month.

former is informed that the feedback is from the AI, while the latter is informed that the feedback is from a human manager. Their performance difference, therefore, results from the AI disclosure effect, holding constant the actual feedback provider as AI. Figure 2 shows that the average performance of the employees in Group 1 (11,211.891 in local currency) is 5.6% lower than that of Group 2 (11,881.887 in local currency). We then compare the employees in Groups 3 and 4, who all receive feedback from human managers; the average performance of Group 3, who are informed that the feedback is from the AI (9,994.130 in local currency), is 4.7% lower than that of Group 4, who are informed that the feedback is from a human manager (10,488.515 in local currency). Hence, regardless of the actual feedback provider, employees who are informed of receiving AI feedback always underperform those are informed of receiving feedback from a human manager. These results thus provide preliminary evidence for H3 on the negative disclosure effect of AI feedback on employee performance (the Online Appendix B shows that the model-based regression results are highly consistent with the model-free evidence presented here).

Deployment Effects of AI Feedback on Employee Job Performance

Note that Figure 2 shows that the magnitudes of the deployment effect measured by Group 1 minus Group 3 and by Group 2 minus Group 4 are similar. Thus, we now examine the average deployment effects by comparing the pooled Groups 1 and 2 (the feedback is generated by the AI for both groups) with the pooled Groups 3 and 4 (the feedback is generated by human managers for both groups). The model is specified in Equation (1) below:

$$Employee\ Performance_i = \alpha + \alpha_1 * Feedback\ Generated\ by\ AI_i + \theta * Controls_i + \varepsilon_{1i} \quad (1)$$

where $Employee\ Performance_i$ is the average payment collected by each employee during the experiment month. The key independent variable is the dummy variable of $Feedback\ Generated\ by\ AI_i$, which is equal to one if *Employee i* receives feedback that is generated by the AI system (i.e., aggregating the employees in Groups 1 and 2), and zero if the employee receives feedback generated by human managers (i.e., aggregating employees in Groups 3 and 4). In addition, $Controls_i$ is a vector of *Employee i*'s characteristics, including prior job performance, age, education level, prior work, tenure at Omega

Corp, and indicators for the managers who provided feedback to them prior to the experiment. Lastly, ε_i is the heteroscedasticity-robust standard error.

Table 2 present the results. The coefficient of $Feedback\ Generated\ by\ AI$ is positive (coeff. = 1,420.120; s.e.= 38.837) in Column (1), suggesting that employees who receive feedback from AI indeed achieve better job performance than those who receive feedback from human managers. We also use alternative dependent variable, by using the log transformation of it to reduce the skewness (Column 2) and using the difference between the performance in the experiment month and that in the previous month as the dependent variable in Column 3. The positive coefficients of $Feedback\ Generated\ by\ AI$ are robust in Columns (2) and (3) (coeff. = 0.132; s.e.= 0.004. and coeff. = 1419.979 s.e.= 38.744, respectively). Figure 3a visualizes the performance difference between the combined Groups 1 and 3 versus the combined Groups 2 and 4. The average collection amount for employees in the groups that actually receive AI feedback is 12.9% higher than the collection amount of employees in the groups that actually receive human feedback (11,559.483 versus 10,241.332 in local currency). Collectively, these results corroborate the positive deployment effect of AI feedback on employee performance, thus supporting H1.

Disclosure Effects of AI Feedback on Employee Job Performance

Figure 2 shows that the difference between the magnitudes of the disclosure effect estimated for Group 1 and Group 2, and that between Group 3 and Group 4 are quite close. Hence, we examine the average disclosure effects by comparing the pooled Groups 1 and 3 (both are informed that they receive AI feedback) with the pooled Groups 2 and 4 (both are informed that they receive human feedback). We estimate the model captured by Equation (2) as follows:

$$Employee\ Performance_i = \beta + \beta_1 * Disclosed\ as\ AI\ Feedback_i + \gamma * Controls_i + \mu_{2i} \quad (2)$$

Here, the key independent variable is $Disclosed\ as\ AI\ Feekback_i$, which equals one if *Employee i* is informed of receiving AI feedback (aggregating all employees in Groups 1 and 3), and zero if the employee is informed of receiving human feedback (grouping employees in Groups 2 and 4).

As reported in Table 2, the negative coefficient of $Disclosed\ as\ AI\ Feedback$ in Column (4) shows that employees who are informed of receiving AI feedback achieve lower job performance than those informed of receiving human feedback (coeff. = -572.803; s.e.= 90.459). The results are robust in Columns (2) and (3) where we use the log transformation of the performance and the difference between the performance in the experiment month and that in the month before as dependent variables, respectively. Figure 3b visually shows that employees who are informed of receiving AI feedback collect 5.4% less payment (10,593.643 in local currency) than those who are informed of receiving human feedback (11,200.683 in local currency). Therefore, these results demonstrate a negative disclosure effect of AI feedback on employee performance, thus supporting H3.

Mechanisms for the Deployment Effect of AI Feedback

To understand why deploying AI feedback may boost employee performance, Omega Corp provides additional data on the feedback content of the experiment, which we use to measure feedback quality and employee learning. Specifically, we construct two objective assessments of the quality of the feedback generated by AI and human managers: “feedback breadth” is the average number of mistakes identified in each feedback email provided to an employee and “feedback depth” is the average number of recommendations made to rectify each identified mistake in the feedback content provided to an employee. We consider higher quality feedback in our data as identifying more mistakes and making more suggestions to correct each mistake. Moreover, we generate an objective measurement of employee learning: *Number of Corrections* captures the number of mistakes identified in the feedback provided during the first week of the experiment that are no longer detected in the feedback emails provided in the fourth week of the experiment for each employee. We consider a larger *Number of Corrections* to reflect a greater extent to which employees learn from the feedback received.

Columns (1) and (2) of Table 3 show that *Feedback Generated by AI* is a positive predictor of the feedback breadth and depth (coeff. = 13.263; s.e.= 0.597 and coeff. = 0.761; s.e.= 0.094, respectively), suggesting that AI feedback points out more mistakes and provides more recommendations to correct each mistake than human managers' feedback (these results are further corroborated by the left panel of Figure A in the online Appendix C, which shows the unconditional mean values of the breadth and depth of AI feedback are indeed greater than those of human feedback). Thus, we conclude that the AI system indeed provides higher quality feedback to employees than the human managers do. As further corroborated by Columns (3), the coefficients of *Feedback Generated by AI* are positive in predicting *Number of Corrections* (coeff. = 10.561; s.e.= 0.446), suggesting that employees indeed learn more from AI feedback than they do from human managers (the left panel of Figure B in the online Appendix C confirms that the unconditional mean value of the number of corrections is higher for employees receiving AI feedback than for those receiving feedback from human managers).

Furthermore, to explore the notion that deploying AI feedback drives job performance through higher quality feedback and greater learning by employees, we employ a mediation analysis using randomized experiment data (Imai et al., 2010) with 1,000 bootstrap replications (Preacher and Hayes 2004). In conducting the mediation analysis, we use *Feedback Generated by AI* as the independent variable, feedback depth, breadth and learning as mediators, and employee performance as the dependent variable; we also include the same control variables.[6] We report all estimation results in Table E1 of the online Appendix E. The results offer some suggestive evidence that relative to human managers, AI provides higher quality feedback, which in turn drives more employee learning and job performance, thus supporting the plausible mechanisms underlying the AI deployment effect in H2.[7]

Mechanisms for the Negative Disclosure Effect of AI Feedback

We examine the mechanisms for the negative disclosure effect in H4 using survey data on employee perceptions. In post-experiment employee surveys, all employees report their perceived trust in

[6] We acknowledge that the mediators are not randomly assigned and thus cannot fully test the causal chain, and the results only provide suggestive evidence that these mediators are relevant explanatory factors.
[7] Agents in call centers have valid concern about potential replacement risk by the AI system because 25% of call center operators are already considering implementing AI system to replace human agents (Bloomberg, 2021).

the quality of the feedback they received from the AI and human managers, as well as the degree of perceived job displacement risk (the specific survey questions are reported in the online Appendix D).

Results in Columns (1) and (2) of Table 4 show that *Disclosed as AI Feedback* is a negative predictor of employees' trust in the quality of feedback (coeff. = -0.784; s.e.= 0.331), and a positive predictor of employees' job displacement risk (coeff. = 2.474; s.e.= 0.151). Thus, disclosing AI feedback induces employees to consider the quality of the feedback to be lower, and heightens their concerns over job replacement risk. Moreover, *Disclosed as AI Feedback* is a negative predictor of the number of corrections in Column (3) of Table 4 (coeff. = -2.781; s.e.= 0.775), suggesting that disclosing the feedback as being generated by AI reduces employees' learning from the feedback. The mediation results summarized in Table E2 of the online Appendix E confirm that disclosing the feedback as being generated by AI (versus human managers) significantly increases employees' negative perceptions in terms of lower trust in the quality of the feedback and higher concerns over job displacement risk, both of which then decrease employees' learning and performance. These results offer suggestive evidence for the underlying mechanisms of the AI disclosure effect in H4.

Note that there exists significant inconsistency because objective metrics show that AI feedback is of higher quality than human feedback (Table 4) but employees' perceptions are the opposite. This inconsistency attests to employees' psychological bias, or aversion to AI algorithm (Leung, Paolacci, and Puntoni 2018; Luo et al. 2019; Mende et al. 2019; Newman, Fast, and Harmon 2020). Furthermore, although employees informed of receiving AI feedback perceive a higher risk of job displacement, this is also inconsistent with Omega Corp's practice of using AI to assist employees to improve their job productivity rather than to replace them. Thus, this might be another form of psychological perceptions against AI.

We also conduct a falsification test. Specifically, we examine whether the feedback breadth and depth might explain the disclosure effect. The coefficients in Columns (4) and (5) of Table 3 are not statistically distinguishable from zero, showing that the objectively measured feedback quality is not different between employees who are informed of receiving feedback from AI and those who are informed of receiving feedback from a human manager (the right panel of Figure A in the online Appendix C reports similar results). These results are expected, because the composition of the *actual* feedback providers among the employees who are informed of receiving AI feedback—the feedback received by about half of them is actually generated by the AI and the feedback for the other half is generated by a human—is the same as that among employees who are informed of receiving human feedback. Thus, our data pass this falsification test.

Employee Tenure Attenuates the Negative Disclosure Effect of AI Feedback

H5 posits that employees' tenure in the company alleviates the negative disclosure effect of AI feedback on job performance. To test this prediction, we first divide employees into four subsample groups based on their tenure with the company, including below the 25th percentile (less than 2 months), between the 25th and 50th percentiles (between 2 and 3 months), between the 50th and 75th percentiles (between 4 and 5 months), and above the 75th percentile (6 months and above). Then we estimate the disclosure effect of AI feedback by rerunning Model (2) in each of the four subsamples; we report the results in Appendix F. The coefficients of *Disclosed as AI Feedback* are negative in Columns (1) and (2) (coeff. = -765.556; s.e.= 313.901 and coeff. = -825.816; s.e.= 417.132, respectively), but they are not statistically different from zero in Columns (3) and (4), confirming that the negative disclosure effect is indeed allayed for employees with a longer tenure in the company. We plot the estimated coefficients of *Disclosed as AI Feedback* in each of the subsample that are divided based on employee tenure in Figure 4. These findings support H5.[8]

[8] New employees often need to spend a nontrivial amount of time to get familiarized with their job responsibility and organizational structure, in which process they start to form their network and social capital in the firm. In call centers, however, this process takes shorter to occur. For example, according to the management team of Omega Corp, agents with 6-month experience are already considered seasoned employees who not only work independently but also start to serve as coaching buddies to newcomers. Moreover, given higher turnover rates of employees in call centers, employees with just a few months of experience are already quite established in the firm. Therefore, even though employee tenure varies only by a few months in the experiment, it continues to create much difference in how established employees are in the company.

Finally, we obtained the break-down data on the performance of each employee during the first 15 days and the second 15 days of the experiment month. Using the panel data and difference-in-differences analysis, we demonstrate how the deployment and disclosure effects unfold over time in Appendix G.

DISCUSSION AND CONCLUSION

Summary of Results

Based on a novel field experiment in a large financial services company, we investigate how using AI to generate feedback on employee performance affects employees' job productivity. First, we demonstrate a positive "deployment effect" that AI feedback, in comparison to human feedback, increases employees' job performance by 12.9%. Moreover, we find that AI provides higher quality feedback in terms of greater breadth and depth than do human managers, which in turn increases employees' learning and performance. Second, we demonstrate a negative "disclosure effect" that the employees who are informed of receiving AI feedback achieve an average performance, which is 5.4% lower than those who are informed of receiving feedback from human managers. We find that employees to whom AI feedback is disclosed tend to have lower trust in the quality of the feedback and higher concerns over job displacement risk, both of which impede their learning and job performance. Furthermore, we show that the value-reducing disclosure effect is less severe among employees who have longer tenure at the firm.

Scope Conditions of Theory

Heterogeneous Managerial Tasks. It is important to note several important scope conditions of this study. We focus on the application of using AI to evaluate employee job performance and provide feedback, including assessment and recommendations, to employees. While these are important managerial functions, they constitute only a small proportion of all tasks that managers need to carry out in managing employees.

More importantly, there exists theoretical distinctions among the managerial tasks in terms of their "structuredness." The managerial functions that we focus on in this study may differ from more sophisticated functions in the degree of structuredness. In our context, while managers and AI need to assess a large amount of unstructured audio data and to comprehend the context of the conversation in generating feedback, one may consider the task of generating codified feedback to be more structured, as there exist quantified performance goals to achieve (to increase collection amount), clear source of information to draw on (phone calls with customers), common understanding of what to look for in the information (to identify mistakes), and shared knowledge of what needs to be done to achieve higher performance ("good" and "bad" scripts). By contrast, other important managerial functions call for managers to tread less-chartered paths where some or all the above assumptions no longer hold. Common examples include (but are not exclusive to) to provide feedback that involve less-codifiable and more-implicit information, to make judgment calls with fewer precedents to follow, to have unstructured conversations regarding promotion and performance improvement opportunities, to communicate and coordinate with team members, and to build personal connections with employees and motivate them.

This critical scope condition generates two important theoretical implications. First, the finding that the deployment of AI feedback outperforms that of human feedback by no means suggests that AI outperforms human managers in performing every managerial task. Instead, it is necessary to investigate whether our analysis can be extended to understanding how AI performs other managerial roles—particularly less structured managerial tasks. For example, while AI excels at making predictions from data and thus may continue to perform well for the managerial tasks that rely on this function, it lacks the abilities to make judgment calls that human managers possess (Miric et al., 2020); hence, it would be an insufficient tool to perform managerial tasks that require judgement calls.

Furthermore, applying AI to less structured managerial functions may provide more sophisticated ways for AI and human managers to create complementarity. Our findings that AI beats human managers in generating higher quality structured feedback (positive "deployment effect" of AI feedback) and human managers beat AI in eliciting favorable perceptions of employees (negative “disclosure effect” of AI feedback) suggest opportunities for AI and human managers to work together to create greater value, such as having human managers communicate to employees the feedback generated by AI. That is, companies can use AI as an effective managerial *assistant* that conducts data analytics and provides feedback content to support human managers’ interactions with employees, thereby keeping humans in the loop. Employees see human managers as their feedback providers, but AI acts as a digital assistant to the managers. Indeed, Jia, Luo, Fang, and Xu (2020) show that human managers who have the transformational leadership style with higher interpersonal skills can more effectively communicate AI-generated feedback to employees, more so than what AI can achieve on its own. Recent research demonstrates more channels for human and AI to complement each other, such as working together to provide more complete inputs into decision-making. For example, Choudhury, Starr, and Agarwal (2020) find that human experts’ domain expertise complements machine learning programs to find prior arts in assessing patents, and Kesavan and Kushwaha (2020) show that retail store managers can use their private information to augment stocking recommendations made by AI algorithms, to achieve higher profit. Therefore, human involvement is not only necessary but may also complement AI’s strength in carrying out more sophisticated managerial functions, such as those occurring in less structured contexts, requiring judgment calls, or benefiting from a high-touch. Thus, complementarity between AI and humans may be created in multiple contexts, albeit through different channels.

Heterogenous Organizational Features. Call centers for loan collection may differ from many other organizations. As discussed earlier when we introduce the background of Omega Corp., persuading customers who are already delinquent to make payments is a tricky task that requires a variety of persuasion skills. Because of these challenges, on-the-job training as embodied in providing performance evaluation and feedback to employees is of critical important to firms in this industry, and employee turnovers are high. As a result, frequent and extensive feedback provision to employees may be more common in this industry than in some other industries, a scope condition that we need to highlight.

Moreover, frequent turnover of employees in this context suggests that employees with longer working experience in the firm may be quite different from newcomers. Although the sample of our study consists of employees who joined the firm relatively recently (but still with notable variation, ranging from 1-7 months), it is fair to caution that the moderating effect of employee tenure on the disclosure effect may be subject to alternative explanations.

Other characteristics of employees may also moderate the disclosure effect. Consistent with the notion that the extent of aversion to an algorithm decreases as human become more familiarity with the algorithm (Kahneman, 2011), employees who are more familiar with AI technologies and their applications in feedback provision may develop greater appreciation of the quality of AI feedback, which may weaken the disclosure effect. Factors that increase employees' familarity with AI technologies may include a younger age, greater exposure to AI either through formal education or other means such as more frequent usage of AI-powered applications. Furthermore, employees who perform well in the firm and thus are given more career development opportunities may have fewer concerns over being replaced by AI technologies, which also can weaken the disclosure effect. Although our empirical context does not afford an opportunity to examine these factors,[9] they are theoretically valuable avenues for future research to explore.

Theoretical Contributions

This study makes several contributions to the academic literature. First, it is imperative to parse out the actual treatment effect of AI deployment from the psychological effects of AI disclosure produced by employees' awareness of the treatment status, and to quantify the degree to which the productivity gain of AI deployment can be offset by the AI disclosure effect. While the overall net effect of using AI

[9] Employees in our sample were all relatively young, ranging between 18 and 22 years old, and had either high school or college education. Thus, there exists insufficient variation in the familiarity with AI technologies that can be proxied by these demographic features. We are unable to obtain the information on the extent to which they use AI-powered applications on their personal electronic devices. We only have information on employees' performance in the month before the experiment but not on their overall performance evaluation or their career opportunities in the firm.

feedback is positive, this estimated net effect is quite an inaccurate measure of the true value of AI because it includes performance loss caused by employees' perceptions. Our back-of-the-envelope calculation shows that on average, the performance loss caused by AI disclosure offsets 43% to 48% of the real value of AI adoption. Thus, without accounting for the disclosure effect created by employees' negative perceptions, the reported positive effect of AI adoption in the literature (mostly developed in economics) substantially under-estimates the true value of AI. Similarly, without accounting for the actual gain of AI adoption, the observed negative effect in the behavioral research created by employees' perceptions against AI is also markedly over-estimated. Scrutinizing the negative AI disclosure effect is important to the literature, because it helps alert firms and prompts them to not only understand some potential psychological reasons (employees lacking trust in the quality of AI feedback and holding greater concerns over the risk of job displacement by AI) but also search for solutions that may mitigate this negative disclosure effect (e.g., employee tenure).

Second, by showing that AI feedback increases employee performance, we reveal a new channel through which new technologies increase firm productivity. While the literature focuses on how AI alters firm's processes of production (Aghion et al., 2017; Aron et al., 2011; Brynjolfsson et al., 2019; Meyer et al., 2014), much less is known about how AI assists with managerial processes such as conducing performance evaluation and feedback for employees. This study thus opens up new avenues for management scholars to go beyond treating AI simply as a factor in the traditional production process toward re-conceptualizing many managerial processes. In doing so, several conventional managerial issues need to be revisited. For example, we show that employees learn more from the higher quality feedback provided by AI than that by human managers, which suggests that AI significantly affects the knowledge transfer within organizations. Furthermore, the finding that disclosing AI feedback triggers negative perceptions among employees that human managers may not have to face suggests that firm management will face new challenges when using AI to manage employees. Thus, despite offering additional opportunities to improve firm management, AI also generates novel issues for firms to resolve (some of which provide promising opportunities for strategy research in the AI era).

Third, this study generates new insights into employees' perceptual bias against being managed by AI. Previous studies theorize that employees question the legitimacy of using AI in management and are concerned about possible infringements on their privacy and autonomy, and a lack of procedural justice (Möhlmann & Zalmanson, 2017b; Newman et al., 2020; Raveendhran & Fast, 2019). Here, we extend this line of work to show that disclosure induces employees' negative perceptions about AI feedback, including lower trust in the quality of the feedback and greater concerns over job displacement risk. Further, we show theoretically and empirically that such perceptions harm employees' actual behavior (learning) and actual performance outcomes (not just perceptions or attitudes), which helps to explain why we need to study employees' aversion to AI in real-world workplaces. Indeed, studying the mechanism of employees' fear of job displacement by AI helps to bridge the "macro"-level research on how AI replaces jobs and reshapes the labor market (e.g., Agrawal, McHale, & Oettl, 2019; Felten et al., 2019; Seamans & Raj, 2018) and the "micro"-level consequences of employees' reactions to AI. These conversations appear in different parts of the literature, but it is crucial to connect them because job displacement risks may generalize negative "spillover" effects by demoralizing employees who do *not* directly face these risks. This issue is important but under-addressed in the extant literature.

Finally, we highlight the heterogeneity in employees' perceptions against AI. This knowledge contributes to the theoretical basis of an emerging but instrumental topic, namely, whether AI complements or substitutes human capital in firms (Choudhury, Starr, & Agarwal, 2020; Fountaine, McCarthy, & Saleh, 2019; Jia et al. 2020). A greater knowledge of who is less susceptible to holding negative perceptions of AI will enable scholars and firms to better understand which employees stand to benefit from and thus complement with the deployment of AI technologies in firms. Therefore, treating employees as a homogenous whole undermines the value of AI applications. By contrast, the knowledge of employee heterogeneity enables scholars and firms to identify subgroups of employees who have greater concerns over AI and address them in a more targeted manner. This approach thus enables firms to create more complementarity and reduce friction in adopting AI technologies at workplace.

Managerial and Policy Implications

For managers, we provide some useful implications. AI performance feedback can be an effective management tool because it reduces time and costs by negating the need to hire human managers to evaluate and train subordinates. Further, AI significantly increases the accuracy and consistency of the analyses of information collected, and generates recommendations that are relevant to each employee and thus help them achieve greater job performance at scale. Because AI feedback enables employees to improve their learning and job performance, all three parties—the firm, employees, and customers served by the firm—may stand to benefit from it.

However, our study also alerts firms to the negative effect of disclosing AI feedback that exists along with the positive impact of deploying AI feedback. We find that employees' negative perceptions offset some of business value of AI feedback, which deserves managerial attention. Further, our finding that the value-destroying disclosure effect is driven by employees' negative perceptions suggests that companies need to be more proactive in communicating with their employees about the objectives, benefits, and scope of AI applications in order to assuage these concerns. Moreover, the result of the allayed negative AI disclosure effect among employees with a longer tenure at the firm suggests that companies may consider deploying AI feedback in a "tiered" instead of a uniform fashion, i.e., using AI to provide performance feedback to veteran employees but using human managers to provide performance feedback novices. The benefit of this strategy is that using AI to provide feedback to employees who are least likely to develop psychological aversion to AI will more fully preserve the productivity gain that AI can generate. However, a potential cost of this strategy may be that employees can try to infer whether different training modes imply that the company values one type more than the other, which might be demoralizing for those who think they are assigned to an "inferior" training group.

If the benefit outweighs the cost, then this strategy will enable firms to achieve even higher returns on AI investment.

For public policymakers, our results offer several implications. With more implementations of AI technologies in businesses, regulations on the transparency of AI usage in the workplace will increase, because AI tilts the power balance in favor of firms against employees (Clarke, 2019; MacCarthy, 2020; National Law Review, 2019). Critics accuse AI applications as being another tool for "[B]osses … seeking to wring every last drop of productivity and labor out of their workers since before computers" (Heaven, 2020). In particular, as more employees work from home (out of sight), managers may keep monitoring them using AI applications (not out of mind). Regulators are concerned that under the veneer of data objectivity, firms might abuse AI to spy on employee and squeeze more value out of them, and even target certain groups for layoffs.

As important as transparency is,[10] does mandating disclosure *alone* help employees protect their wellbeing? Our study shows that disclosing AI feedback leads employees to trust the quality of the feedback to a lesser degree (although AI feedback is of higher quality) and to perceive more job displacement risk (despite the fact that the goal of deploying AI is to increase employee performance instead of replacing employees), both of which are negative perceptions that reduce employees' learning and performance. These outcomes generate a deadweight loss that only adds to the psychological burdens on employees without benefiting any stakeholder, and thus they reduce the value "pie" for all to share. Therefore, the designs of public policies on using AI in managing firms need to be more holistic. A direct implication is that while transparency is pivotal, the mandate of disclosing AI applications needs to be complemented by other policy instruments that directly tackle employees' doubts over AI. These can be achieved by multiple means, including public discourse, education, and more importantly, systematic

[10] Transparency may exceed simple disclosure of the act of using algorithms, to include transparency in the mechanism, the purpose, and the data used to train the AI algorithms.

support for retraining human talents to handle higher-skill innovative tasks, while AI assists humans in lower-skill repetitive tasks.

Future Research

Will the negative disclosure effect wane or even dissipate over time? As employees and society become more familiar with and learn more about the value of AI application in business operations, this may become a real possibility. Kahneman (2011) indeed postulates that algorithm aversion may be weakened with greater familiarity with the use of algorithm, and Logg, Minson, & Moore (2019) discuss that more common use of algorithm increases the acceptance of algorithmic advice such as in the case of weather forecasts. Therefore, greater familiarity with how AI feedback functions, more recognition of AI's potential value in generating higher quality feedback, and a better understanding of how AI feedback can be used to help rather than replace employees may all gradually alleviate employees' aversion to AI, thereby reducing the negative disclosure effect.

Furthermore, might the disclosure of AI applications even generate positive perceptions among employees and increase their work motivation? It is theoretically plausible because if employees are educated and convinced that the application of AI feedback acts as a form of organizational support offered by their firms to help them improve job performance and career opportunities in the future, then they might be better motivated and increase their job performance. A change of context could also reshape the disclosure effect. For example, in contexts with a concrete, external standard of accuracy for investment decisions or sports predictions, people rely more heavily on advice given by the disclosed algorithm than that by a human expert (Logg et al, 2019). These theoretical possibilities deserve closer examination by the future research.

REFERENCES

Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: implications of technology for growth, factor shares, and employment. *American Economic Review*, *108*(6), 1488–1542. https://doi.org/10.1257/aer.20160696

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, *128*(6), 2188–2244. https://doi.org/10.1086/705716

Aghion, P., Jones, B., & Jones, C. (2017). *Artificial Intelligence and Economic Growth*. https://doi.org/10.3386/w23928

Agrawal, A, Gans, J., & Goldfarb, A. (2018). *Prediction machines: the simple economics of artificial intelligence*. Harvard Business Press.

Agrawal, Ajay, Gans, J., & Goldfarb, A. (2016). Managing the Machines: AI is making prediction cheap, posing new challenges for managers. *Working Paper*, *October*, 1–14. https://static1.squarespace.com/static/528e51b6e4b0234f427a14fb/t/581a32e6d482e9494ba441c0/1478111975274/EconomicsOfAI.pdf

Agrawal, Ajay, Gans, J. S., & Goldfarb, A. (2019). Artificial intelligence: The ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives, 33*(2), 31–50. https://doi.org/10.1257/jep.33.2.31

Agrawal, Ajay, McHale, J., & Oettl, A. (2019). *Artificial Intelligence, Scientific Discovery, and Commercial Innovation*. http://www.esri.go.jp/jp/workshop/190730/esri2019_second_presenter_paper.pdf

Aron, R., Dutta, S., Janakiraman, R., & Pathak, P. A. (2011). The impact of automation of systems on medical errors: Evidence from field research. *Information Systems Research*, *22*(3), 429–446. https://doi.org/10.1287/isre.1110.0350

Ashforth, B. (1994). Petty Tyranny in Organizations. *Human Relations*, *47*(7), 755–778. https://doi.org/10.1177/001872679404700701

Bloomberg. (2021). *Artificial Intelligence, Chatbots Threaten Call-Center Industry, Human Operators - Bloomberg*. Bloomberg.Com. https://www.bloomberg.com/news/articles/2021-03-16/artificial-intelligence-chatbots-threaten-call-center-industry-human-operators?mod=djemAIPro

Brynjolfsson, E., Hui, X., & Liu, M. (2019). Does machine translation affect international trade? Evidence from a large digital platform. *Management Science*, *65*(12), 5449–5460. https://doi.org/10.1287/mnsc.2019.3388

Bughin, J., & Manyika, J. (2019). *Your AI Efforts Won't Succeed Unless They Benefit Employees*. Harvard Business Review. https://hbr.org/2019/07/your-ai-efforts-wont-succeed-unless-they-benefit-employees

Carpenter, R. (2019). *Advancements in AI and Its Impact on Human Employees*. HR Technologist. https://www.hrtechnologist.com/articles/digital-transformation/advancements-in-ai-and-its-impact-on-human-employees/

Cheatham, B., Javanmardian, K., & Samandari, H. (2019). *Confronting AI risks* . McKinsey Quarterly. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/confronting-the-risks-of-artificial-intelligence#

Choudhury, P., Starr, E., & Agarwal, R. (2020). Machine learning and human capital complementarities: Experimental evidence on bias mitigation. *Strategic Management Journal*, *41*(8), 1381–1411. https://doi.org/10.1002/smj.3152

Clarke, Y. D. (2019). *H.R.2231 - 116th Congress (2019-2020): Algorithmic Accountability Act of 2019*. https://www.congress.gov/bill/116th-congress/house-bill/2231

Colangelo, M. (2020). *Mass Adoption Of AI In Financial Services Expected Within Two Years*. Forbes. https://www.forbes.com/sites/cognitiveworld/2020/02/20/mass-adoption-of-ai-in-financial-services-expected-within-two-years/#58e29b667d71

Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business*

Review.
Deloitte. (2019). *Global perspectives on AI | Deloitte Insights*. https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/global-perspectives-ai-adoption.html
Dietvorst, B., Simmons, J. P., & Massey, C. (2015). Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err. *Journal of Experimental Psychology: General*, *144*(1), 114–126. https://doi.org/10.1037/xge0000033
Ewert, A. (1984). Employee Resistance to Computer Technology. *Journal of Physical Education, Recreation & Dance*, *55*(4), 34–36. https://doi.org/10.1080/07303084.1984.10629723
Felten, E., Raj, M., & Seamans, R. C. (2019). The Effect of Artificial Intelligence on Human Labor: An Ability-Based Approach. *Academy of Management Proceedings*, *2019*(1), 15784. https://doi.org/10.5465/ambpp.2019.140
Fleming, N. (2018). How artificial intelligence is changing drug discovery. *Nature*, *557*(7707), S55–S57. https://doi.org/10.1038/d41586-018-05267-x
Fountaine, T., McCarthy, B., & Saleh, T. (2019). *Building the AI-Powered Organization*. Harvard Business Review. https://hbr.org/2019/07/building-the-ai-powered-organization
Garimella, K. (2018). *Job Loss From AI? There's More To Fear!* Forbes. https://www.forbes.com/sites/cognitiveworld/2018/08/07/job-loss-from-ai-theres-more-to-fear/#123daab923eb
Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*. https://doi.org/10.5465/annals.2018.0057
Goyena, R., & Fallis, A. . (2019). Strategy and Strategic Alignment of Professional Service Firms. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004
Heaven, W. D. (2020). *This startup is using AI to give workers a "productivity score" | MIT Technology Review*. MIT Technology Review. https://www.technologyreview.com/2020/06/04/1002671/startup-ai-workers-productivity-score-bias-machine-learning-business-covid/?mod=djemAIPro
Huang, M. H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459
Hunt, J. W., & Saul, P. N. (1975). The Relationship of Age, Tenure, and Job Satisfaction in Males and Females. *Academy of Management Journal*, *18*(4), 690–702. https://doi.org/10.5465/255372
Iansiti, M., & Lakhani, K. R. (2020). Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World. Harvard Business Press
Imai, K., Keele, L., & Tingley, D. (2010). A General Approach to Causal Mediation Analysis. *Psychological Methods*, *15*(4), 309–334. https://doi.org/10.1037/a0020761
Ip, G. (2019). *For Lower-Paid Workers, the Robot Overlords Have Arrived - WSJ*. The Wall Street Journal. https://www.wsj.com/articles/for-lower-paid-workers-the-robot-overlords-have-arrived-11556719323
Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577–586. https://doi.org/10.1016/J.BUSHOR.2018.03.007
Jia, N., Luo, X., Fang, Z., & Xu, B. (2020). *Can Artificial Intelligence Substitute or Complement Managers? Divergent Outcomes for Transformational and Transactional Managers in a Field Experiment, Working Paper*.
Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science (New York, N.Y.)*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415
Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.
Kesavan, S., & Kushwaha, T. (2020). Field experiment on the profit implications of merchants' discretionary power to override data-driven decision-making tools. *Management Science*, *66*(11), 5182–5190. https://doi.org/10.1287/mnsc.2020.3743

Latham, G. P., & Kinne, S. B. (1974). Improving job performance through training in goal setting. *Journal of Applied Psychology*, *59*(2), 187–191. https://doi.org/10.1037/h0036530

Lee, K.-F. (2018). *AI Superpowers: China, Silicon Valley, and the New World Order*. Houghton Mifflin Harcourt.

Leung, E., Paolacci, G., & Puntoni, S. (2018). Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior. *Journal of Marketing Research*, *55*(6), 818–831. https://doi.org/10.1177/0022243718818423

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*(151), 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019a). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019b). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013

Luo X, Qin S, Fang Z, Qu Z (2021), "Artificial Intelligence Coach for Sales Agents: Caveats and Solutions," Journal of Marketing, 85 (March), 14-32. DOI: 10.1177/0022242920956676.

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, *38*(6), 937–947. https://doi.org/10.1287/mksc.2019.1192

MacCarthy, M. (2020). *AI needs more regulation, not less*. Brookings Institution. https://www.brookings.edu/research/ai-needs-more-regulation-not-less/

Marr, B. (2018). *The Amazing Ways How Unilever Uses Artificial Intelligence To Recruit & Train Thousands Of Employees*. Forbes. https://www.forbes.com/sites/bernardmarr/2018/12/14/the-amazing-ways-how-unilever-uses-artificial-intelligence-to-recruit-train-thousands-of-employees/?sh=23d2eabb6274

Martinez, A. (2019). *Considering AI In Hiring? As Its Use Grows, So Do The Legal Implications For Employers.* Forbes. https://www.forbes.com/sites/alonzomartinez/2019/12/05/considering-ai-in-hiring-as-its-use-grows-so-do-the-legal-implications-for-employers/#195258fe77d4

Mende, M., Scott, M. L., van Doorn, J., Grewal, D., & Shanks, I. (2019). Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses. *Journal of Marketing Research*, *56*(4), 535–556. https://doi.org/10.1177/0022243718822827

Meyer, G., Adomavicius, G., Johnson, P. E., Elidrisi, M., Rush, W. A., Sperl-Hillen, J. A. M., & O'Connor, P. J. (2014). A machine learning approach to improving dynamic decision making. *Information Systems Research*, *25*(2), 239–263. https://doi.org/10.1287/isre.2014.0513

Mintzberg, H. (1990). The design school: Reconsidering the basic premises of strategic management. *Strategic Management Journal*, *11*(3), 171–195. https://doi.org/10.1002/SMJ.4250110302

Miric, M., Lu, J., & Teodoridis, F. (2020). Decision-Making Skills in an AI World: Lessons from Online Chess. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3538840

Möhlmann, M., & Zalmanson, L. (2017a). *Navigating Algorithmic Management and Drivers*. https://www.researchgate.net/publication/319965259

Möhlmann, M., & Zalmanson, L. (2017b). Navigating Algorithmic Management and Drivers. In *researchgate.net*. https://www.researchgate.net/publication/319965259

National Law Review. (2019). *Keeping an Eye on Artificial Intelligence Regulation and Legislation*. Natlawreview.Com. https://www.natlawreview.com/article/keeping-eye-artificial-intelligence-regulation-and-legislation

Newman, D., Fast, N., & Harmon, D. (2020). When Eliminating Bias Isn't Fair: Algorithmic Reductionism and Procedural Justice in Human Resource Decisions. *Organizational Behavior and Human Decision Processes*.

O'keefe, J., Moss, D., Martinez, T., & Rose, P. (2019). *Professional Perspective AI Regulation and Risks to Employers*. http://bna.com/copyright-permission-request/

Oldham, G. R., & Cummings, A. (1996). Employee Creativity: Personal and Contextual Factors at Work. *Academy of Management Journal*, *39*(3), 607–634. https://doi.org/10.5465/256657

Perry, R. W., & Mankin, L. D. (2004). Understanding Employee Trust in Management: Conceptual Clarification and Correlates. *Public Personnel Management*, *33*(3), 277–290. https://doi.org/10.1177/009102600403300303

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731. https://doi.org/10.3758/BF03206553

Premuzic, T. C., Wade, M., & Jordan, J. (2018). *As AI Makes More Decisions, the Nature of Leadership Will Change*. Harvard Business Review. https://hbr.org/2018/01/as-ai-makes-more-decisions-the-nature-of-leadership-will-change

Raveendhran, R., & Fast, N. (2019). Humans Judge, Technologies Nudge: When and Why People Embrace Behavior Tracking Products. *Academy of Management Proceedings*, *2019*(1), 13103. https://doi.org/10.5465/ambpp.2019.13103abstract

Roe, D. (2018). *How AI Can Negatively Impact Employee Experiences*. CMS Wire. https://www.cmswire.com/digital-workplace/how-ai-can-negatively-impact-employee-experiences/

Rogan, M., & Mors, M. L. (2014). A Network Perspective on Individual-Level Ambidexterity in Organizations. *Organization Science*, *25*(6), 1860–1877. https://doi.org/10.1287/orsc.2014.0901

Roose, K. (2019). *A Machine May Not Take Your Job, but One Could Become Your Boss - The New York Times*. The New York Times. https://www.nytimes.com/2019/06/23/technology/artificial-intelligence-ai-workplace.html

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.

Schanke, S., Butch, G., & Ray, G. (2020). Estimating the Economic Impact of ‘ Humanizing ’ Customer Service Chatbots. *Working Paper*, 1–35. https://osf.io/preprints/socarxiv/xud6z/

Seamans, R., & Raj, M. (2018). AI, Labor, Productivity, and the Need for Firm-Level Data. *National Bureau of Economic Research*. https://doi.org/10.3386/w24239

Sun, C., Shi, Z. (June), Liu, X., Ghose, A., Li, X., & Xiong, F. (2019). The Effect of Voice AI on Consumer Purchase and Search Behavior. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3480877

Tarafdar, M., Beath, C. M., & Ross, J. W. (2019). Using AI to enhance business operations. *MIT Sloan Management Review*, *60*(4), 37–44.

Taylor, F. (1911). *Scientific management*. https://sites.suffolk.edu/govt521/files/2010/09/taylor.pdf

Teich, D. (2019). *AI Chatbots: Companies Love Them; Consumers, Not So Much..*. Forbes. https://www.forbes.com/sites/davidteich/2019/04/09/ai-chatbots-companies-love-them-consumers-not-so-much/#74bac054284c

Thiel, W. (2019). *The role of Artificial Intelligence in customer experience*. Www.Pointillist.Com. https://www.pointillist.com/blog/role-of-ai-in-customer-experience/

Verma, J. P., & Agrawal, S. (2016). Big Data Analytics: Challenges And Applications For Text, Audio, Video, And Social Media Data. *International Journal on Soft Computing, Artificial Intelligence and Applications*, *5*(1), 41–51. https://doi.org/10.5121/ijscai.2016.5105

Webb, M., Autor, D., Bloom, N., Bresnahan, T., Brynjolfsson, E., Chetty, R., Coyle, D., Gentzkow, M., Hoxby, C., Jaravel, X., Jones, C., Klenow, P., Pistaferri, L., Rafey, W., Sorkin, I., Reenen, J. Van, Agrawal, A., Connelly, T., Han, A., … Thornton, G. (2019). The Impact of Artificial Intelligence on the Labor Market. In *papers.ssrn.com*. https://web.stanford.edu/

Webster, J. (1993). Turning work into play: Implications for microcomputer software training. *Journal of Management*, *19*(1), 127–146. https://doi.org/10.1016/0149-2063(93)90049-s

Whitt, W. (2006). Staffing a call center with uncertain arrival rate and absenteeism. *Productins and Operations Management*, *15*(1), 88–102. https://doi.org/10.7916/D8M3377T

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of

Educational Feedback Research. *Frontiers in Psychology*, *10*, 3087. https://doi.org/10.3389/fpsyg.2019.03087

Table 1: Summary Statistics and Randomization Check

Panel A	*Prior Job Performance*	*Age*	*Education*	*Prior Call Center Working*	*Tenure*
Mean	9540.796	20.351	1.418	0.075	3.218
Standard Deviation	1397.036	1.315	0.494	0.265	1.336
Minimum	6679	18	1	0	1
Maximum	12148	22	2	1	7

Note: *Prior Job Performance* is the average daily collection amount achieved by the agent in the month before the experiment. *Education* is a categorical variable: 1=High school degree, 2=College degree; *Prior Call Center Working* is a binary variable with 1=the agent has working experience in a call center before joining the company, and 0=otherwise; *Tenure* is the length of time the employee has worked in the company by month

Panel B	N	*Prior Job Performance*	*Age*	*Education*	*Prior Call Center Working*	*Tenure*
Feedback Generated by & Feedback Provider Disclosed as AI	64	9400.578	20.422	1.422	0.062	3.047
Feedback Generated by AI but Feedback Provider Disclosed as Human Managers	69	9564.986	20.689	1.478	0.043	3.130
Feedback Generated by Human Managers but Feedback Provider Disclosed as AI	66	9613.742	20.511	1.485	0.091	3.273
Feedback Generated by and Disclosed as Human Managers	66	9578.530	20.089	1.288	0.106	3.424
F-value/ Chi Square		0.30	1.74	2.30	2.277	1.01
P-value		0.829	0.1596	0.077	0.517	0.389

Panel C	N	*Job Performance*	*Feedback Breadth*	*Feedback Depth*	*Number of Corrections*
Feedback Generated by and Feedback Provider Disclosed as AI	64	1811.312	23.469	1.891	14.109
		(16.590)	(1.037)	(0.158)	(0.619)
Feedback Generated by AI but Feedback Provider Disclosed as Human Managers	69	2316.901	24.522	1.928	18.435
		(31.268)	(0.317)	(0.056)	(0.302)
Feedback Generated by Human Managers but Feedback Provider Disclosed as AI	66	380.388	11.015	1.167	5.258
		(8.185)	(0.448)	(0.055)	(0.232)
Feedback Generated by and Feedback Provider Disclosed as Human Managers	66	909.985	10.742	1.167	6.258
		(21.486)	(0.353)	(0.051)	(0.340)

Note: *Job Performance* is the average daily collection amount achieved by the employee during the 30-day experiment period. *Feedback Breadth* is the average number of mistakes identified in each feedback email provided to an employee. *Feedback Depth* the average number of recommendations made to rectify each identified mistake in the feedback email provided to an employee, and *Number of Corrections* is he number of mistakes identified in the feedback provided during the first week of the experiment that are no longer detected in the feedback emails provided in the fourth week of the experiment for each employee.

Table 2: Effects of Deploying and Disclosing AI Feedback on Employee Performance

	(1)	(2)	(3)	(4)	(5)	(6)
	Deployment Effect of AI Feedback on Employee Performance			Disclosure Effect of AI Feedback on Employee Performance		
Dependent Variable	*Job Performance*	*Job Performance*	*Job Performance (Difference)*	*Job Performance*	*Job Performance*	*Job Performance (Difference)*
Model	OLS	Log OLS	OLS	OLS	Log OLS	OLS
Feedback Generated by AI	1420.120	0.132	1419.979			
	(38.837)	(0.004)	(38.744)			
Disclosed as AI Feedback				-527.803	-0.048	-525.866
				(90.459)	(0.008)	(90.285)
Prior Job Performance	1.000	0.000		0.980	0.000	
	(0.014)	(0.000)		(0.032)	0.000	
Age	7.131	0.001	7.190	79.847	0.008	79.315
	(30.260)	(0.003)	(30.266)	(73.651)	(0.007)	(73.534)
Education	-18.674	-0.003	-18.137	-66.979	-0.008	-73.441
	(80.997)	(0.008)	(80.453)	(193.556)	(0.018)	(193.323)
Prior Working	-22.399	0.000	-22.426	-251.111	-0.021	-251.109
	(60.936)	(0.006)	(60.981)	(148.515)	(0.014)	(149.257)
Tenure	21.048	0.003	21.241	-25.947	-0.002	-28.149
	(14.423)	(0.001)	(14.309)	(33.316)	(0.003)	(33.107)
Indicators of Pre-experiment Managers	Y	Y	Y	Y	Y	Y
Constant	403.706	8.306	417.695	326.939	8.299	165.015
	(552.361)	(0.053)	(529.642)	(1305.541)	(0.122)	(1284.863)
N	265	265	265	265	265	265
R-Squared	0.962	0.960	0.848	0.795	0.792	0.163

Standard errors are reported in the parentheses.

Table 3: Mechanisms of the Positive Impact of Deploying AI Feedback

	(1)	(2)	(3)	(4)	(5)
Dependent Variable	*Feedback Breadth*	*Feedback Depth*	*Number of Corrections*	*Feedback Breadth*	*Feedback Depth*
Model	OLS	OLS	OLS	OLS	OLS
Feedback Generated by AI	13.263	0.761	10.561		
	(0.597)	(0.094)	(0.446)		
Disclosed as AI Feedback				-0.437	-0.027
				(1.027)	(0.103)
Prior Job Performance	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age	-1.151	-0.105	-0.541	-0.368	-0.060
	(0.490)	(0.082)	(0.369)	(0.852)	(0.090)
Education	2.872	0.283	1.469	1.886	0.227
	(1.305)	(0.218)	(1.004)	(2.160)	(0.241)
Prior Working	-1.345	-0.320	-1.329	-3.562	-0.447
	(1.197)	(0.124)	(0.832)	(1.336)	(0.101)
Tenure	0.354	0.063	0.029	-0.009	0.042
	(0.226)	(0.034)	(0.159)	(0.387)	(0.039)
Indicators of Pre-experiment Managers	Y	Y	Y	Y	Y
Constant	25.494	2.865	12.210	20.058	2.555
	(9.029)	(1.434)	(6.759)	(15.216)	(1.570)
N	265	265	265	265	265
R-Squared	0.665	0.235	0.698	0.020	0.029

Standard errors are reported in the parentheses.

Table 4: Mechanisms of the Negative Effects of Disclosing AI Feedback

	(1)	(2)	(3)	(4)	(5)
Dependent Variable	*Trust in Feedback Quality*	*Perceived Job Replacement Risk*	*Number of Corrections*	*Trust in Feedback Quality*	*Perceived Job Replacement Risk*
Model	OLS	OLS	OLS	OLS	OLS
Disclosed as AI Feedback	-0.784	2.474	-2.781		
	(0.331)	(0.151)	(0.775)		
Feedback Generated by AI				4.697	-0.913
				(0.177)	(0.214)
Prior Job Performance	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.029	0.161	0.026	-0.292	0.158
	(0.273)	(0.116)	(0.639)	(0.144)	(0.180)
Education	0.334	-0.615	0.974	0.609	-0.391
	(0.693)	(0.318)	(1.666)	(0.371)	(0.465)
Prior Working	-0.992	-0.103	-3.050	-0.218	-0.212
	(0.645)	(0.297)	(1.044)	(0.287)	(0.386)
Tenure	-0.115	0.088	-0.301	0.024	0.021
	(0.127)	(0.051)	(0.293)	(0.064)	(0.080)
Indicators of Pre-experiment Managers	Y	Y	Y	Y	Y
Constant	6.370	0.040	10.438	7.634	2.232
	(4.842)	(2.064)	(11.282)	(2.549)	(3.274)
N	265	265	265	265	265
R-Squared	0.067	0.521	0.092	0.753	0.090

Standard errors are reported in the parentheses.

Figure 1: Field Experimental Design

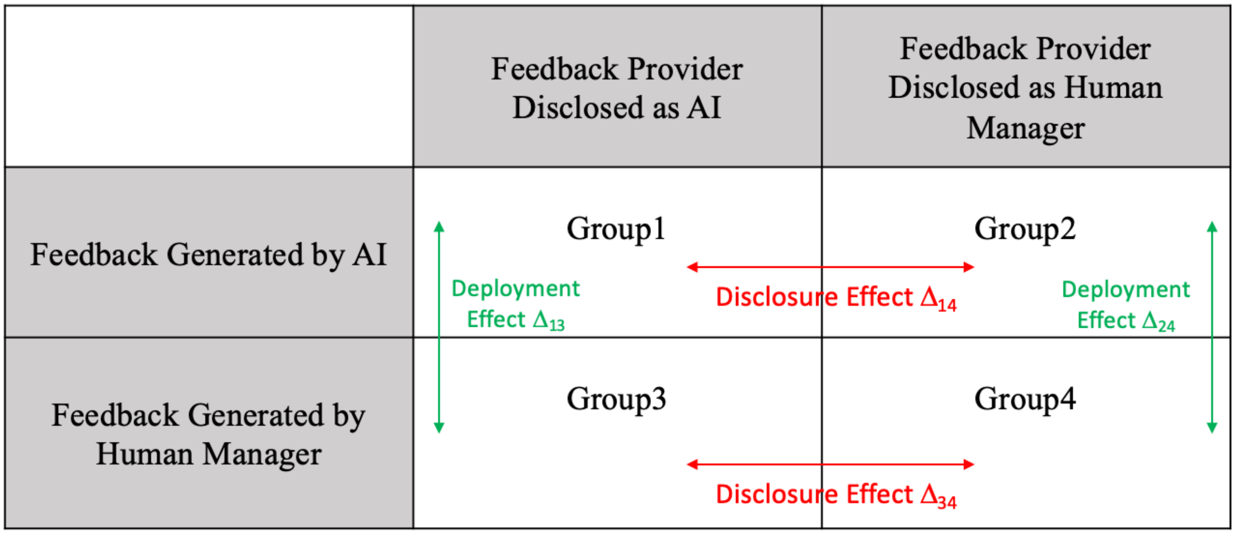

Figure 2: Performance Comparison among the Four Experimental Conditions

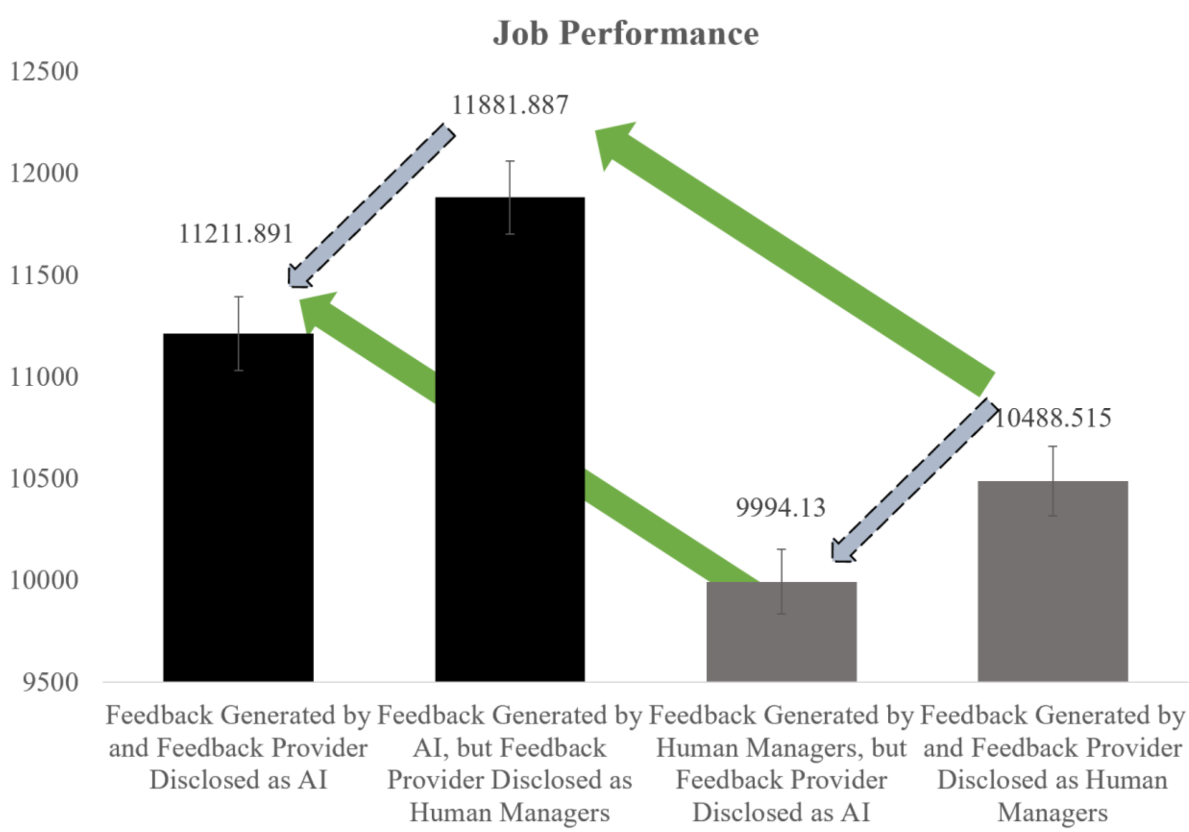

Figure 3a: Deployment Effect of AI Feedback (Unconditional mean comparison)

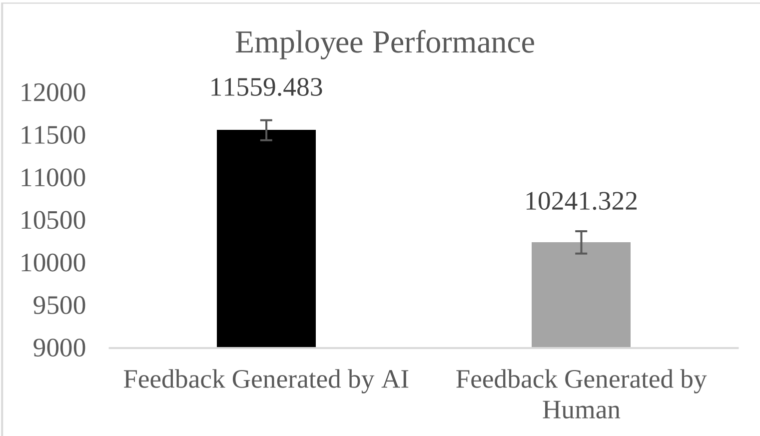

Figure 3b: Disclosure Effects of AI Feedback (Unconditional mean comparison)

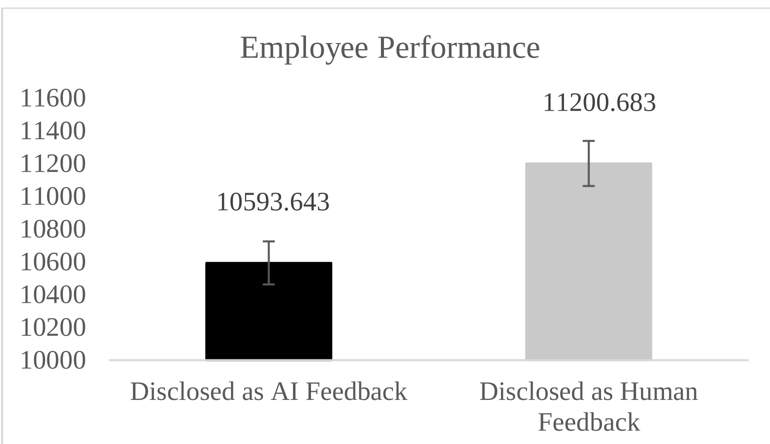

Figure 4. Disclosure Effect of AI Feedback in Subsamples of Employees with Varying Tenure