Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Performance Effects of Giving Front-Line Employees Direct Access to Performance Data and Thereby Limiting the Supervisor's Feedback-Intermediation Role: Evidence from a Field Experiment

Ethan Bernstein, Shelley Xin Li

To cite this article:
Ethan Bernstein, Shelley Xin Li (2026) The Performance Effects of Giving Front-Line Employees Direct Access to Performance Data and Thereby Limiting the Supervisor's Feedback-Intermediation Role: Evidence from a Field Experiment. Management Science 72(2):805-835. https://doi.org/10.1287/mnsc.2022.02395

https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 72, No. 2, February 2026, pp. 805–835
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

The Performance Effects of Giving Front-Line Employees Direct Access to Performance Data and Thereby Limiting the Supervisor's Feedback-Intermediation Role: Evidence from a Field Experiment

Ethan Bernstein,[a,*] **Shelley Xin Li**[b,*]

[a] Organizational Behavior, Harvard Business School, Boston, Massachusetts 02465; [b] Accounting, USC Marshall School of Business, Los Angeles, California 90089

*Corresponding authors

Contact: e@hbs.edu, https://orcid.org/0000-0001-9819-0639 (EB); shelley.li@marshall.usc.edu, https://orcid.org/0000-0001-9811-8098 (SXL)

Received: August 6, 2022
Revised: November 23, 2023; June 1, 2024
Accepted: June 26, 2024
Published Online in Articles in Advance: May 15, 2025

https://doi.org/10.1287/mnsc.2022.02395

Abstract. This paper examines how giving front-line employees direct access to their performance data affects performance. To explore that impact, we conducted a field experiment at a service organization that made employees' daily time-use analytics—previously available only to supervisors—simultaneously available to the employees themselves. We find, compared with the preintervention mean value, a significant treatment effect (an 11% decrease) in nonproductive time relative to the control group. That time, however, flows not strictly to productive (revenue-generating) activities but largely to the most convenient outlets, suggesting—as supported by our qualitative evidence—that data transparency on average shifted behavior more toward avoiding nonproductive activities than toward approaching productive activities. (As one participant observed, it led people to "conform, not excel.") We examine three relational factors we believed, based on prior feedback research, could moderate the performance effect: perceived supervisor support, social comparison orientation, and work motivation type. Performance improvements are greater for employees who perceived their supervisors as less supportive and for those with low intrinsic motivation or high extrinsic motivation; we fail to find a moderating effect of social comparison orientation. Therefore, although we identify the avoidance (although not approach) value of transparent performance data, our results also tell a nuanced story about the supervisor's optimal role in delivering feedback: The ability to shift people away from "bad" and toward other uses of time depends on the employee's perception of supervisor quality and the employee's motivation type.

History: Accepted by David Simchi-Levi, organizations.

Funding: We are grateful to the Division of Research and Faculty Development at the Harvard Business School and the USC Marshall School of Business for funding this research.

Supplemental Material: The data files are available at https://doi.org/10.1287/mnsc.2022.02395.

Keywords: organizational studies: effectiveness-performance • organizational studies: information • organizational studies: personnel • organizational studies: productivity • transparency • performance feedback • front-line workers • human resource management • information sharing • management control systems • direct feedback • disintermediation

1. Introduction

Performance feedback—information about the effectiveness of individuals' work behaviors (Ashford and Cummings 1983, Taylor et al. 1984)—has long been an important tool for managing front-line performance (Ilgen et al. 1979, Balcazar et al. 1985, Larson 1989, Alvero et al. 2001, Lourenço 2016). It gives organizations a concrete way to address a common, often vexing, employee question: How am I doing?

Traditionally, supervisors have been responsible for answering that question—and for doing so well. A long-standing literature investigates how they should aggregate, filter, frame, and then deliver performance feedback to maximize its positive effects (Kluger and DeNisi 1996). However, advances in technology that allow for greater quantification of employee behaviors (Pierce et al. 2015, Mazmanian and Beckman 2018, Ranganathan and Benson 2020) and for more immediate reporting of them (Mollick and Rothbard 2013, Goler et al. 2016, Staats et al. 2017) are changing supervisors' role in the feedback process. In settings where employees can directly access their performance data (for example, via digital

scorecards that show results for individual members of a workgroup), they need not rely solely on their supervisor for the information.

In this paper, we theorize that phenomenon—direct employee access to performance data previously accessed only by supervisors—as partial disintermediation of the supervisors. We say "partial" because, although employees can access the information without a supervisor's involvement, they may still receive it through supervisor feedback or seek a supervisor's help in utilizing the information (Steelman and Wolfeld 2018). In other, nonmanagerial contexts, scholars have used the term *apomediation* to describe such partial disintermediation of "middle(wo)men," upon observing how the Internet allows consumers (like patients in healthcare settings) to directly access information that was previously exclusive to expert intermediaries (like their doctors) (Eysenbach 2008). We argue that something similar is happening in management to supervisors delivering feedback at work. In prior scholarship on nonmanagerial contexts, it has been said that an intermediary "stands in between (Latin: inter- means in between) the consumer and information, meaning that she is a necessary mediating agent to receive the information in the first place," whereas an apomediary will instead "stand by (Latin: apo- means separate, detached, away from) to guide a consumer to high quality of information and services without being a prerequisite to obtain that information or service in the first place and with limited individual power to alter or select the information that is being brokered" (Eysenbach 2008, p. 6). Likewise, in the workplace, supervisors who serve as intermediaries stand between employees and their performance data and can thus control the messaging around it, whereas those who serve as apomediaries stand by to facilitate understanding and use while also allowing employees simultaneous direct access to the analytics. In our paper, we build on this terminology to describe an emerging approach to management in workplaces with increasingly transparent performance data and reporting to front-line employees.

Yet even as more organizations give employees direct access to performance data (Cunningham and McGregor 2015, Cappelli and Tavis 2016, Ewenstein et al. 2016) and the supervisor role therefore becomes more apomediary than intermediary, empirical research has only begun to investigate the impact of that shift on performance. In the management literature's long tradition of measuring and reporting on tasks and output—harkening back to the time-motion studies of Taylor (1914) and Gilbreth and Gilbreth (1919)—only a handful of empirical studies (notably Earley (1988) and Kluger and Adler (1993)) have investigated the implications of computer-delivered versus human-delivered employee feedback (Table A.1). Although those older studies examined situations in which the same, simplistic feedback (e.g., correct/incorrect) was delivered by either a computer or a human, the workplace has since evolved, thanks in no small part to a proliferation of new technological tools and capabilities (Leonardi and Neeley 2022). For that reason, it is now time to look at how performance can be affected in real work settings when the same raw—but in this century, more complex—data that used to be accessible only by supervisors (who aggregated, filtered, framed, and delivered it to employees as feedback) becomes directly accessible by employees.

That is what we set out to do in this paper. In particular, we examine how front-line performance responds to direct access to performance data—and, if different employees respond differently, who benefits from directly accessed feedback (with supervisor as apomediary) versus supervisor-delivered feedback (with supervisor as intermediary, exclusively accessing the raw data).

On the one hand, technologists have good reasons to anticipate a positive potential impact: As data generation and algorithmic evaluation capabilities have advanced (Madhavan and Wiegmann 2007, Lee et al. 2015, Kellogg et al. 2020, Cameron 2022), scholars have found that giving employees direct access to performance data can ward off certain problems associated with supervisor-intermediated feedback (Frick and Simmons 2008, Jackson 2012, Schyns and Schilling 2013, Scott 2017). For example, it can mitigate the risk that supervisors will put a subjective "spin" on the results (Prendergast 1999, Gibbs et al. 2004, Bol 2008, Harvey and Green 2022) or that their feedback may be skewed by personal biases (Larson 1989, Adams 2005, Botelho and Gertsberg 2022), inconsistencies, or supervisor-employee polarization (Patil and Bernstein 2022). In addition, direct access to the data makes employees less beholden to a supervisor, thus allowing them to work around a lower-quality supervisor or seek advice from others, including peers, when they know their performance is subpar (Ashford and Cummings 1983, Larson 1989). Finally, direct data access may promote constructive changes in employee behavior. Much as finance and supply chain scholars have found positive effects on transactional and logistical performance from the partial disintermediation of "middlemen" through transparent data (Chircu and Kauffman 1999), management scholars might expect gains in front-line performance due to employees' sharper focus on the task-motivation and task-learning components of their own cognitive processes, which the established feedback intervention theory (FIT; Kluger and DeNisi 1996) predicts will lead to higher performance.

On the other hand, directly accessing the data can also harm employees' performance. Supervisors traditionally have served as feedback intermediaries because they can do things as gatekeepers that data cannot do alone (Kraut et al. 1989). Prior research in organizational behavior suggests that—like intermediaries in other

fields (Chircu and Kauffman 1999, Spulber 1999)—supervisors aggregate, filter, and frame performance data in valuable ways. For instance, they can minimize information-processing costs for employees, assist in interpreting the data, focus employees' attention on the most relevant information, and offer actionable guidance (Alvero et al. 2001, London 2003). In contrast, when performance data flow directly to employees without supervisor intermediation, people may have greater difficulty finding, recognizing, making sense of, and using what's important (Cases-Arce et al. 2017), much as healthcare consumers without any doctor mediation may "'get lost' in the vast amount of information [online] and arrive at the wrong or irrelevant information" (Eysenbach 2008, p. 6). Without help building a narrative (Hersel 2020) around the performance data, especially in contexts where overall performance is more subjective than what can be captured by just the numbers, employees may find direct data access more confusing than useful.

Alternatively, a third option is possible: Direct data access might have little impact on employee performance at all—or might affect it only for certain kinds of employees. Those outcomes would also be interesting, as they might suggest that the managerial effort being devoted to making these data transparent is not a performance revolution (Cappelli and Tavis 2016) but rather (at least in some situations) a technology-driven but not performance-enhancing shift in organizational processes.

We use a field experiment with embedded-participant observation at a large U.S. gas utility (GasCo) to empirically investigate how the simple act of providing front-line employees with transparent access to their own and other team members' performance data—the exact same data previously accessed only by supervisors—affects their performance. Prior to our study, the research site was representative of many business organizations in terms of the differences in ex ante information access between supervisors and front-line employees: The supervisors alone could see three daily performance metrics, focused on effective time use,[1] for all employees under their supervision—automatically tracked by sensors in employees' trucks, computers, and handheld devices[2]—and could then use those analytics to provide performance feedback. During our study, a randomly selected group of front-line employees ("service mechanics") received direct ("transparent") access to individual-level performance for those three metrics. These were *exactly* the same data that had previously been available only to their supervisors and that supervisors generally shared with employees during their interactions each day. Over the following months, we compared this treatment group's performance with that of the control group, which did not receive direct access to the data.

We combine three empirical approaches to estimate the effect on front-line employee performance of partially disintermediating supervisors—that is, casting them in the role of apomediary—with transparent performance data. First, we use a traditional difference-in-differences estimation, where we compare employee performance in four treated work centers, before and after they received the performance data previously reserved exclusively for their supervisors, with employee performance in seven untreated (control) work centers. Our second empirical strategy is to collect and analyze survey data to investigate several key moderating effects that both deepen our understanding of why partial disintermediation affects performance and mitigate concerns about alternative explanations. Our third approach uses embedded participant observation to complement our main results and provide qualitative evidence about mechanisms.

Our baseline findings point to the value of transparent performance data, relative to supervisor-intermediated performance feedback, for influencing front-line employees, *on average*, to reduce time-use labeled "nonproductive" but not to directly focus on increasing "productive" time use (with nonproductive and productive having been clearly defined, prior to our study, jointly by the organization and its employees). Specifically, compared with the preintervention mean value, the treatment group experienced an 11% decrease in nonproductive time relative to the control group, but on average, we see immaterial effects (in both magnitude and significance) on productive time and on a third category of time, support time (non–revenue-generating work activities). Those results suggest that employees—when given an opportunity to make sense of performance data without supervisor intermediation—focused more on avoiding activities perceived as bad than on pursuing, or approaching, activities perceived as good.

Two moderating effects allow us to unpack this baseline result and explore the heterogenous effects across different types of employees, allowing us to disentangle *for whom* managers-as-feedback-givers added value and for whom managers-as-feedback-givers got in the way. First, we find that employees who perceived low supervisor support ex ante had treatment effects of greater magnitude and statistical significance, consistent with a view of transparent performance data as a partial substitute for lower-quality supervisors in the delivery of feedback. Second, we find that employees with lower intrinsic motivation and/or higher extrinsic motivation—in other words, employees more driven by external rather than internal rewards—also had treatment effects of greater magnitude and statistical significance, consistent with a view of transparent performance data as a vehicle for providing public recognition and, thus, external performance incentives.

Shifting supervisors from intermediaries to apomediaries might increase employees' access to relative performance information (Kuhnen and Tymula 2012, Tafkov 2013), given that supervisors might be more hesitant than computers to share information about other employees' performance. We therefore explored whether the performance improvements were driven by greater access to relative performance information. We find little evidence of that. If the main result was largely driven by that access, we would expect greater treatment effects for employees with stronger social comparison orientation, as prior work has shown (Hannan et al. 2013). But we find mixed, inconclusive evidence: Those with *weak* social comparison orientation in fact reduced nonproductive time more than they increased productive time, whereas those with strong social comparison orientation increased productive time more than they reduced nonproductive time, although the latter effect is statistically insignificant. Based on the moderator analyses, our main effect appears not to be driven simply by greater access to relative performance information.

Our results were reinforced by the qualitative evidence collected by our embedded participant observer. Employees appreciated seeing detailed performance data whenever they wanted. But they spoke much less about improving their productive time than about not wanting to "stand out" for their nonproductive time. Compared with other employees, those dissatisfied with their supervisors' level of support welcomed access to transparent performance data with more enthusiasm and perceived greater benefit from it.

This paper has implications for several research streams. First, we contribute to research on performance feedback, both building on FIT and providing new nuance to it. Although a vast literature on the effects of performance feedback has clarified the relationship between performance improvement and certain feedback attributes (such as frequency, specificity, and negative or positive framing) or certain recipient attributes (such as gender and past performance), and whether and how the feedback is linked to incentives or targets (Hannan et al. 2008, Thornock 2016, Eyring and Narayanan 2018, Lourenço et al. 2018), none directly examines the effect of eliminating the information-transparency gap between employees and supervisors in a real work setting with complex measures of performance. Our study—by specifically evaluating the effect of giving employees direct access to raw performance data and thereby partially disintermediating managers in the feedback process—empirically underscores the dual importance of transparent data *and* high-quality apomediaries (supervisors) in performance feedback, whereas also showing a downside to making supervisors apomediaries.

Second, this paper contributes to growing work on the interaction between information systems and human behavior. Specifically, consistent with an older research stream on how "bad is stronger than good" (Baumeister et al. 2001, p. 323), greater data transparency shifted attention toward avoiding behaviors perceived to be "bad" (nonproductive) rather than toward approaching behaviors perceived to be "good" (productive) or neutral (support). It also activated or even amplified extrinsic (but not intrinsic) motivation. These findings add to prior findings on human-computer versus human-human interaction. By removing the additional layer of filtering that a human intermediary (such as a supervisor) provides, human-computer interactions may reduce "customization" in the type of information disclosed or in the matching of feedback to individual motivation types. As a net result, making supervisors apomediaries may privilege employees' avoidance motives over employees' approach motives (Roth and Cohen 1986, Elliot and Church 1997).

Finally, this paper contributes to emerging research on novel organizational structures, as enabled by technologies that permit the partial disintermediation of middle managers (Tapscott and Tapscott 2017, Bernstein 2022, Reitzig 2022). For instance, scholars have identified information provision as one of the key roles of traditional managers (Puranam et al. 2014). Although that function may soon be taken over by technology, at this point we know little about the effects on how work is organized and done. Those merit investigation. As for the impact on outcomes, whereas the intermediation literature has not investigated supervisors as internal mediaries, our results are consistent with previous findings on disintermediation, apomediation, and reintermediation, in very different domains and contexts. Our study both reinforces the importance of considering mediary quality in assessing the performance effects of partial disintermediation (Chircu and Kauffman 1999) and extends this research into fresh territory by examining end users' attributes (such as motivation types) as moderators. Furthermore, broadening conceptualizations of "disintermediation" and "apomediation" may open up other fruitful avenues for organizational research. For example, as researchers continue to investigate the influence of technology-enabled transparency and quantification (Ranganathan and Benson 2020) on operations management and learning systems (Staats 2018), our findings can help redefine—and reshape—the evolving role of the manager (Mintzberg 1975, 2013).

2. Motivation and Theory

Studies in psychology, organizations, and operations demonstrate some beneficial effects of feedback interventions on front-line performance (Ilgen et al. 1979, Pritchard et al. 1981, Ivancevich and McMahon 1982). However, empirical research also demonstrates that such feedback does not always improve performance

(Eriksson et al. 2009, Lourenço 2016); in fact, it can undermine performance for the best employees (Haas and Hayes 2006), for the worst employees (Podsakoff and Farh 1989), or for anyone in between (Ivancevich and McMahon 1982, Kluger et al. 1994).

In the most notable and highly cited attempt to bring together such mixed findings, the review of empirical studies of Kluger and DeNisi (1996), accounting for 607 effect sizes and 23,663 observations, found that, although feedback interventions did improve performance on average, 38% of them made it worse. To explain the heterogenous effects, Kluger and DeNisi (1996) used a meta-analysis of those studies to develop the FIT, which argues that the variance between positive and negative performance outcomes stems from where the feedback focuses an employee's attention: on task-learning processes (how a task is done), on task-motivation processes (how much effort goes into a task), or on meta-task processes (how the employee interprets the feedback and the implications of that interpretation for affect, framing, and the self):

> Specifically, a feedback intervention (FI) provided for a familiar task, containing cues that support learning, attracting attention to feedback-standard discrepancies at the task level (velocity FI and goal setting), and void of cues to the meta-task level (e.g., cues that direct attention to the self) is likely to yield impressive gains in performance, possibly exceeding one standard deviation (Kluger and DeNisi 1996, p. 278).

In short, they conclude that cues directing employee attention to meta-task processes make feedback less effective, whereas those directing attention to task-motivation or task-learning processes make it more effective. Given those findings, it is worth exploring which types of cues supervisors have been providing in their role as feedback givers.

2.1. What Is the Supervisor's Role in the Delivery of Performance Feedback?

In the nearly 30 years since FIT was published, supervisors (direct managers) have, in practice, traditionally played a central role in crafting feedback that could meet FIT's propositions for front-line employees. We conceptualize such supervisors as intermediaries—people who act as a link between others (Oxford English Dictionary 2015)—because their feedback-crafting functions mirror the three functions that scholars have distilled from the activities of a wide variety of intermediaries (e.g., go-betweens, middlemen, brokers, traders, diplomats, bridgers, mediators, recruiters, wholesalers, spies, interpreters, priests, rabbis) (Smith 2002, McCubbrey and Taylor 2005, Howells 2006, Hurt 2007, Zhang and Li 2010, Ahn et al. 2011, Chiappa et al. 2014, Stanton and Thomas 2015): aggregating, filtering, and framing (Spulber 1999).

Like intermediaries in other domains,[3] supervisors typically receive more information about their supervisees than those supervisees do and must then decide how to aggregate, filter, and frame that data and provide periodic feedback. They may aggregate real-time or hourly performance data to provide daily, weekly, monthly, or quarterly feedback. They may filter the data to focus on what they consider most relevant to a particular employee (for example, filtering out factors outside the employee's control). Supervisors can also frame the data to indicate good or bad performance based on certain benchmarks. It might then seem to follow that good supervisors—for example, those who deliver performance feedback as FIT suggests it is most optimally delivered—would, like good intermediaries in other domains, drive the best outcomes with their aggregating, filtering, and framing. That has been the traditional view of many organizations seeking better performance from feedback interventions. But organizations are changing with technology. Workplace data are increasingly transparent to employees, and the supervisor's role is evolving along with that shift.

2.2. How Might Performance Data Transparency—and Its Partial Disintermediation of Supervisors—Affect Front-Line Outcomes?

Just as traditional middlemen in other domains are disintermediated by technology in one or more aspects of their roles (Chircu and Kauffman 1999), a similar trend is happening among supervisors in the delivery of performance feedback to front-line employees. Some organizations are choosing to make the same performance data, previously available only to supervisors, transparent to employees through real-time tracking, dashboards, and other analytics tools. Akin to the often-cited observation by Eysenbach (2008) that a patient's ability to directly access information on the Internet has shifted doctors and nurses from traditional *inter*mediaries (gatekeepers of information who stand in between the patient and the information) to *apo*mediaries (agents who stand by to potentially guide a patient to high quality of information without being gatekeepers to it), transparent performance data turn supervisors into apomediaries in their role of delivering performance feedback. How will that affect employee performance?

Prior empirical studies—although scarce, disparate, and sometimes conflicting—offer some foundations for a theoretical answer. For starters, two studies in the 1980s and 1990s—which focused only on the delivery of the same simple information via either a computer or a person—explored the impact of computer-mediated versus human-mediated feedback. They found that (a) computer-generated feedback to mail order processors led to more trust, stronger feelings of self-efficacy, and better performance than did identical feedback from a human supervisor (Earley 1988), and (b) people seemed

to be more likely to seek binary outcome feedback (i.e., correct versus incorrect) from a computer than from a human (Kluger and Adler 1993) because interacting with a machine involved less stigma, less impression management, and more frequent access (Ashford and Cummings 1983, Karabenick and Knapp 1988). Given such promising early results, Kluger and Adler (1993, p. 2) encouraged more work on this topic when they wrote that there had been "much discussion concerning the relative merits of person- versus computer-mediated performance feedback, but little empirical research"—a sentiment unfortunately still reiterated by Stone et al. (2015) and Johnson et al. (2015) 22 years later. Although scholars have built on the initial two studies in attempts to optimize computer-mediated feedback (see Table A.1), particularly in the fields of education (Neri et al. 2008, Murphy 2010) and alcoholics counseling (Elliott et al. 2008, Carey et al. 2012, Cole et al. 2018), we are aware of no empirical work that extends this research to contemporary organizational contexts, with more complex metrics (versus binary correct/incorrect or yes/no feedback on reaching a simple goal), more intricate and professional (Leicht and Fennell 2001) work (versus simple tasks), and more sophisticated tracking (versus simple output-based metrics).

That lack of progress may stem, in part, from the fact that after FIT was published, practitioners have primarily interpreted FIT as a means to improve the delivery of performance feedback *by supervisors*. After all, supervisors—like other intermediaries who have survived and even thrived despite technological attempts to partially disintermediate them (Chircu and Kauffman 1999, Jallat and Capek 2001, Fang et al. 2015)—add various kinds of value when delivering performance feedback. That value could easily get lost in organizations that bypass parts of the traditional feedback process by giving employees direct access to the same performance data supervisors receive. For example, as is argued in many cases of disintermediation (Chircu and Kauffman 1999), people with direct access may struggle to find or recognize the most relevant information or to realize what actions would be most appropriate, leading to distraction or confusion. In short, as in other domains where disintermediation has become technologically more feasible (Spulber 1999), the intermediary—in this case, the supervisor delivering performance feedback to the front-line employee—may provide performance-enhancing value, above and beyond the value of the raw data, by aggregating, filtering, and framing that data prior to providing feedback.

However, that is not the only way to interpret FIT. Indeed, one key insight—that higher velocity task-level data will improve performance by directing employees' attention to task learning and task motivation—might conversely support a prediction of better performance when employees are given direct access to performance data previously available only to their supervisors. A second key insight of FIT—that reducing "cues to the meta-task level" will improve performance—might similarly reinforce the merits of data transparency; that is, employees with direct data access may be less likely to feel evaluative pressure and to distract themselves from performance improvement by thinking about what the supervisor's view of their performance means for them and their self-esteem.

Even as organizations choose to provide direct employee access to performance data, our field's ability to predict the consequences remains limited, therefore, by a lack of empirical work directly comparing the performance effects of supervisor-intermediated data with those of direct data access in contexts involving complex feedback (e.g., not just correct/incorrect), more professional work (not just basic routine tasks), and advanced technology (capable of automated, faster, more sophisticated analytics). This uncertainty frames our main research question: how will the act of providing front-line employees with direct access to transparent performance data (the same data previously intermediated by supervisors) affect performance outcomes?

2.3. What Factors Would Influence the Performance Effects of Transparent Performance Data?

Given our main research question above and our read of the performance feedback literature, we wondered whether different employees might respond differently to direct data access versus supervisor-mediated feedback and, if so, what might explain their varied responses (in particular, their behavioral shifts toward good or bad uses of time (Podsakoff and Farh 1989, McFarland and Miller 1994, Van Dijk and Kluger 2011)). Because data transparency might encourage individual employees to reallocate effort and time among different types of tasks (Brewer 1995, Hannan et al. 2013), we wanted to look not just at average effects within and across task types but also at *any individual shift* in time use.

Why study time use versus other performance metrics? As contemporary front-line work has become more complex (Bordoloi et al. 2019)—and therefore more "professional" (Leicht and Fennell 2001)—than the simpler routines studied in the feedback literature in the past (Earley 1988, Kluger and Adler 1993), supervising an employee's judgment about which task to do by tracking their allocation of time has become more tractable than supervising how a complex task is done and the outcome it produces. As a result, employers are increasingly adopting systems that track and codify front-line workers' time use, often breaking that time down into three categories of tasks: billable/productive tasks (such as customer-facing or other revenue-generating activities), nonbillable/nonproductive tasks

(such as personal phone calls and coffee breaks), and support/investment tasks (necessary work activities, such as training and maintenance, that do not directly generate revenue but make long-term revenue generation possible) (Maister 1993, Evans et al. 2004). Numerous organizations from a range of industries—such as healthcare (Kc and Terwiesch 2009, Kippenbrock et al. 2018), construction (Guidry et al. 2018), drilling (York et al. 2009, Eren 2018), trucking (Roberti et al. 2014), manufacturing (Berman et al. 1997), law (Campbell and Charlesworth 2012), creative industries (Pitts 2022), and financial services (Staats and Gino 2012)—have found that scheduling (Pinedo 2012), tracking, and supervising time allocation is at least as important as supervising how a task is ultimately performed.

Employees then often receive feedback on how they allocate time, just as they receive feedback on how they do tasks. Organizations—across the various industries above—are experimenting with giving employees direct access to data on how well they are using their time rather than leaving it to supervisors to aggregate, filter, frame, and then deliver that feedback. How that transparency affects time allocation remains unknown, however, especially as prior feedback research has focused more on how tasks are completed than on the types of tasks employees choose to spend their time doing.

Yet, judgments about employees' time allocation remain intact. Some task types are deemed good investments of time and others bad. Because research has shown many ways in which "bad is stronger than good" (Baumeister et al. 2001, p. 323), and because research has shown that feedback valence affects which behaviors are encouraged (Peifer et al. 2020), it might be that making individual performance transparent across multiple categories of tasks—without supervisor intermediation—would naturally focus employee attention more on avoiding activities that could be considered bad (nonproductive) than on increasing activities considered to be good (productive). Alternatively, research on using leaderboards and gamified dashboards at work to motivate performance suggests that competition for the top spots naturally focuses employee attention more on the most positively viewed activities (Yakura 2001, Chanen 2005, Mollick and Rothbard 2013, Bernstein and Blunden 2015)—an approach rather than avoidance motive. There is even theoretical support for transparent performance data driving greater support time, as employees see others investing in themselves now to excel in the future (Carpenter et al. 2012). We thus investigate task type—productive/billable, nonproductive/unbillable, or support/investment—for baseline effects of giving employees direct access to data that were previously used only by supervisors to provide feedback.

As with some prior empirical research on performance feedback, the most interesting findings may emerge from investigating potential moderators of the performance effects (Kluger and DeNisi 1996, Alvero et al. 2001, Barends et al. 2022). Theory on workplace transparency (Bernstein 2017) suggests that when performance data are made more transparent, three mutually exclusive categories of *relational* moderators—relational because they incorporate effects attributed to the employee's relationship with surrounding others (Eberly et al. 2011, 2017)—may influence how an individual pays attention and reacts to the transparent data. The three categories are (1) vertical relational moderators (based on a hierarchical relationship), such as how supervisor or direct report dynamics influence the individual's attention and behaviors; (2) horizontal relational moderators (based on a peer relationship), such as the degree to which dynamics among teammates influence the individual's attention and behaviors; and (3) relational-versus-self moderators (based on the relative influence of external relationships or internal motivations), such as the degree to which external/internal forces influence the individual's attention behaviors. In each case, because previously private data are now more visible to others, how the individual responds can be affected by those relational attributes.

We therefore also study three critical relational moderators, one from each category above, that existing theory suggests would be most relevant for our study: (1) perceived supervisor support (as a vertical relational moderator), because employees' response to raw data versus supervisor-mediated feedback might depend on their view of supervisor quality (Snyder et al. 1984, Feys et al. 2008); (2) *social comparison orientation* (as a horizontal relational moderator), because employees' response to raw data versus supervisor-mediated feedback might depend on their sensitivity to relative performance information (Blanes et al. 2011, Song et al. 2018); and (3) *work motivation type* (as a relational-versus-self moderator), because employees' response to raw data versus supervisor-mediated feedback might depend on the extent to which they are motivated by external pressure (Kuvaas 2006, DePasque and Tricomi 2015).

2.3.1. Perceived Supervisor Support: Quality of Intermediary (Vertical Relational Moderator). As in other studies of disintermediation (Fang et al. 2015), the quality of the supervisor—the "value" that he or she is perceived to deliver—may moderate the effects of partial disintermediation. In our investigation, how employee performance shifts when supervisors no longer hold information privilege and are no longer exclusively relied on for feedback would seem to depend on how much employees value the intermediary's support above and beyond the delivery of raw data. Perceived supervisor support has been tied to in-role and extra-role performance (Eisenberger et al. 2002, Shanock and Eisenberger 2006, Jokisaari and Nurmi 2009), but to our

knowledge, that connection's impact on the performance effects of transparent data remains unstudied. Based on existing work, we theorize two potential opposing outcomes.

On the one hand, access to the same performance data that supervisors receive can only convey *what* one's performance is and not *how* to improve it (Hellervik et al. 1992, London and Smither 1995). Supervisor feedback is more than just data; the supervisor can also provide guidance. The volume and complexity of "raw" performance data could even heighten the need for a high-quality intermediary who can curate, simplify, and interpret relevant information. Seifert et al. (2003) found that feedback without a facilitator (i.e., just a report) failed to improve performance, whereas the same feedback with a facilitator was perceived as more useful and did indeed improve performance. Good supervisors are often conceptualized as "coaches" (Gilley and Gilley 2007) who excel at noninformational dimensions of the role, such as providing guidance and inspiration, that would *complement* transparent performance data. Giving employees direct access to the data previously reserved for the supervisor would, in effect, encourage more—and higher value—conversation with the supervisor. In that case, the larger performance gain from access to such data could come from those who work under higher-quality supervisors—those perceived to be more supportive.

On the other hand, access to transparent data may be a *substitute* for part of the supervisor's role in performance feedback (Tapscott and Ticoll 2003, Hamel 2011), especially for "bad" supervisors whom employees do not view as supportive. In those cases, employees may value the data source more than the middleman: Finding out where they stand relative to coworkers can provide them with other role models—peers—who may be better coaches than their own supervisors (Ilgen et al. 1981, Larson 1986), making the supervisor's input even less valuable than it had been while also boosting performance.

2.3.2. Social Comparison Orientation: Effects of Relative Performance Information (Horizontal Relational Moderator). The seminal paper of Festinger'(1954) proposed that, in the absence of clear standards of correctness, people evaluate themselves, their opinions, and their capabilities in comparison with others (Wood 1989, Gibbons and Buunk 1999, Suls et al. 2002). Knowing how one performs relative to peers can improve performance (Azmat and Iriberri 2010, Kuhnen and Tymula 2012, Tafkov 2013). Although the "desire to learn about the self through comparison with others is universal," the extent to which people do so has been shown to vary, as measured by the Iowa-Netherlands Comparison Orientation Measure (Gibbons and Buunk 1999, p. 199). Partially disintermediating supervisors with transparent performance data may increase employees' access to relative performance information if supervisors previously chose to buffer the disclosure of such information, viewing it as sensitive. For our purposes, if greater access to relative performance information (Song et al. 2018)—rather than supervisor disintermediation—is driving outcomes, we would expect that those more prone to social comparison would be more likely to improve after gaining such access. Therefore, we test for that moderating effect.

2.3.3. Work Motivation Type: External Pressure via Greater Transparency (Relational-vs.-Self Moderator). A vibrant—and sometimes heated—dialogue has emerged about the impact of transparent data on motivation in the workplace (Christin 2018, Aarons-Mele 2020). If viewing transparent performance data (e.g., a digital dashboard showing daily performance metrics for all employees in the same work unit) generates higher external performance pressure than the less-transparent supervisor-led feedback process, whether it yields positive or negative performance effects would depend on employees' type and level of motivation, as has been true in much empirical research on feedback (Anderson and Rodin 1989, Kluger and DeNisi 1996).

Affected by personality types (Swift and Peterson 2018), employees are motivated for different reasons (Amabile et al. 1994): some by a passionate interest in or deep enjoyment of what they are doing (i.e., high intrinsic motivation) and others by external inducements such as pay or social recognition (i.e., high extrinsic motivation). Although prior empirical evidence is lacking, it may be that intrinsically motivated employees are less affected by such pressure from transparent data because they are less sensitive to external pressures that are filtered out, or at least buffered, by an intermediating supervisor, whereas extrinsically motivated front-line employees are more affected. In other words, transparent performance data could act as a complement to extrinsic motivation and thus as a substitute for intrinsic motivation (for employees who lack it). Conversely, the opposite may be true: Because the data offers an opportunity for additional learning (typically an intrinsic motivator) but not for additional rewards (in the absence of financial or other extrinsic incentives), transparent performance data could act as a complement to intrinsic motivation and thus as a substitute for extrinsic motivation (for employees who lack it). We test for both possibilities.

3. Methods and Data

3.1. Research Setting

The context of our study is a service operation—a natural gas distribution company (referred to as GasCo, a

pseudonym) serving approximately 425,000 customers in the southeastern United States. Most of GasCo's 1,100 employees are customer facing, including the cadre of field-based professional service technicians, known as mechanics, on whom this study focuses.

Mechanics spend their days on the road addressing customer requests to turn on or turn off gas, repair gas appliances, repair leaks, and respond to emergencies. They typically start their day by logging into the system on their trucks and then reviewing and accepting orders made available by dispatch. An algorithm automatically maps a path to an order, after which the mechanic drives there, arrives on site, completes the order, and then drives to the next order or task. That continues throughout the day, with the exception of one to two blocks of time per day when they typically meet back at the work center with their supervisor and other mechanics. Although the activities may appear routine, mechanics self-identify as and are considered highly trained professionals because of (a) the risk inherent in any activity involving gas, (b) their substantial training, and (c) the wide variability in the contexts, systems, and devices they are expected to safely diagnose and fix. Their activities meet our two activity-based criteria for this study: sufficiently specified for performance metrics to be comparable across individuals but sufficiently complex to permit wide variation in results based on capability and on factors largely within the individual worker's control.

The mechanics' context also meets two criteria for our study. First, they interact with customers onsite and sometimes with other mechanics in their own work centers (based on geographic areas), but rarely with mechanics in other centers. Randomization of the experimental intervention at the work-center level was therefore unlikely to suffer from "contamination" (i.e., mechanics in the control group were unlikely to learn about the difference in treatment status through interactions with those in the treatment group). Second, and equally important, GasCo's workforce did not face the high-powered economic incentives (positive or negative) that are far more prevalent in prior experiments investigating the effectiveness of performance feedback (Holmstrom and Milgrom 1991, Larkin et al. 2012) than they are in a real workplace. In fact, many, if not most, front-line jobs lack strong financial incentives for performance (see U.S. Bureau of Labor Statistics Employee Costs for Employee Compensation survey, summarized in Gittleman and Pierce (2013, table 1)). GasCo's employee incentive plan was relatively disconnected from individual performance: Monetary incentives were based on companywide objectives, and even if the company met or exceeded those, the maximum incentive bonus payment for an individual employee was 2.5% of total compensation. Nor were there strong career-related incentives: Because the workforce was unionized (all mechanics in the study were union members, and one of them, M15, was a union leader), there was less concern about job loss based on individual performance, and it was understood that promotions were based primarily on tenure. This context allowed us to observe the effect of direct access to performance data *itself*, decoupled from financial incentives or fear of career consequences.

Not long before our study, GasCo had consolidated the customer service organizations, which included mechanics, of all its acquisitions. The mechanics now worked from 11 work centers, the staffing of which ranged from 2 to 42 mechanics. Because the consolidation involved integrating previously autonomous organizations with different histories, the performance feedback systems also needed to be integrated into a consistent set of metrics. Through a bottom-up effort (including input from the mechanics and their unions), GasCo had generated a single scorecard of time-use data—collected automatically by technology in the mechanics' trucks and computers (not self-reported) and thus both minimally subject to gaming and effortlessly measured—to which the mechanics had collectively agreed. All these data rolled-up to three umbrella categories of performance metrics: mechanics' percentages of productive time, support time, and nonproductive time. (These are defined in the Dependent Variables section below.)

3.2. Data Collection and Measures

3.2.1. Field Experiment. Four of the 11 work centers were randomly selected for the treatment condition, which involved being able to access—using any computer (including mechanics' truck laptops) and an individual's company account login—an intranet page with a scorecard displaying the performance metrics of all mechanics in the same work center. These were the same data that supervisors previously received (and continued to receive). The other seven work centers served as a control group operating at the status quo: the performance data were available to a supervisor, who then delivered feedback to each employee. Thirty-one mechanics were thus randomly assigned to the treatment group and 92 to the control group.[4] The only difference between the two experimental conditions was whether the same performance data were transparent to all (in the treatment group, where both supervisors and mechanics could see the data) or available only to supervisors (in the control group, where mechanics received performance information only via supervisor intermediation). According to our embedded participant-observer, daily feedback conversations during work center meetings continued in both treatment and control conditions, so the only change was that the mechanics had direct access to the raw data in the treatment group.[5]

This intervention was implemented as a natural field experiment, which allowed us to draw causal inferences

(due to the random selection of the treatment work centers) and to examine the effects of access to transparent performance data in a natural context. Subjects were unaware that they were participating in a study, allowing us to avoid self-selection and discard alternative explanations such as the Hawthorne effect (Mayo 1933, Roethlisberger and Dickson 1939).

Figure 1 shows the randomization outcome, treatment conditions, and timeline of the intervention. Figure 2 shows a sample screenshot of the daily scorecard information visible (with names disguised) to individual mechanics in the treatment group—the same performance data that supervisors alone had previously received. Mechanics in the treatment group received an email every morning with a link to the scorecard information. To ensure the quality of the intervention—that is, to be sure mechanics were actually accessing the performance data—we tracked how often the intranet webpages were accessed, although, due to both technological and human subject limitations, we could not identify who had made a particular visit to the scorecard. The experimental intervention ("pilot") ran from June 25 to August 29. We retrieved daily performance data from GasCo's archive for June 1 of the prior year through August 29, the final day of our field experiment; that is, 389 days before the intervention and 65 days during it. We selected a 12-month preintervention period to take into account the full cycle of seasonal effects.

To gather moderating and control variables, we sent a Qualtrics survey link by email on June 19 (six days before the intervention). Mechanics could access it from any computer. We made clear that the survey was conducted by external researchers, not GasCo, and that no responses would be seen by anyone in the company. We sent the survey to all mechanics and had a 51% response rate, which management reported was typical for this population. We compare the key employee

Figure 2. (Color online) Sample Screenshot of the Mechanic's View of the Daily Scorecard Information and Time/Activity Categorization

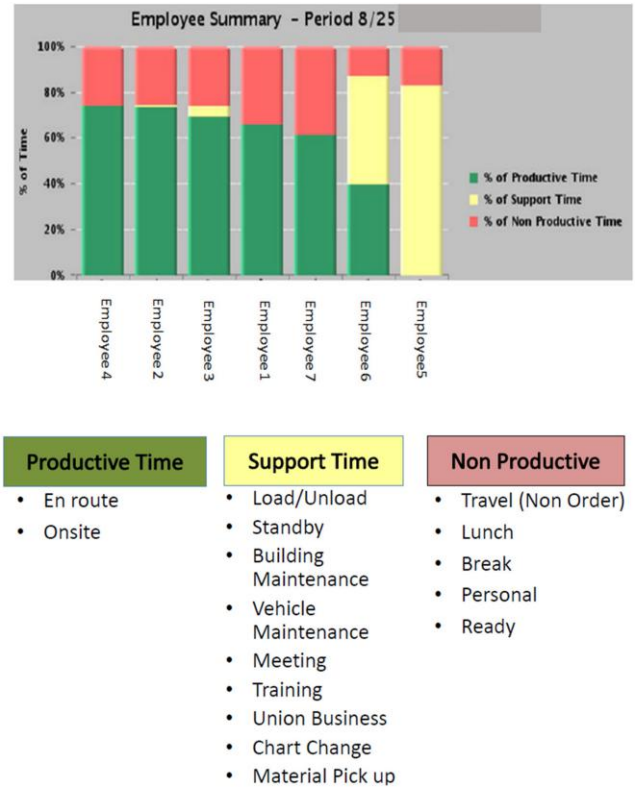
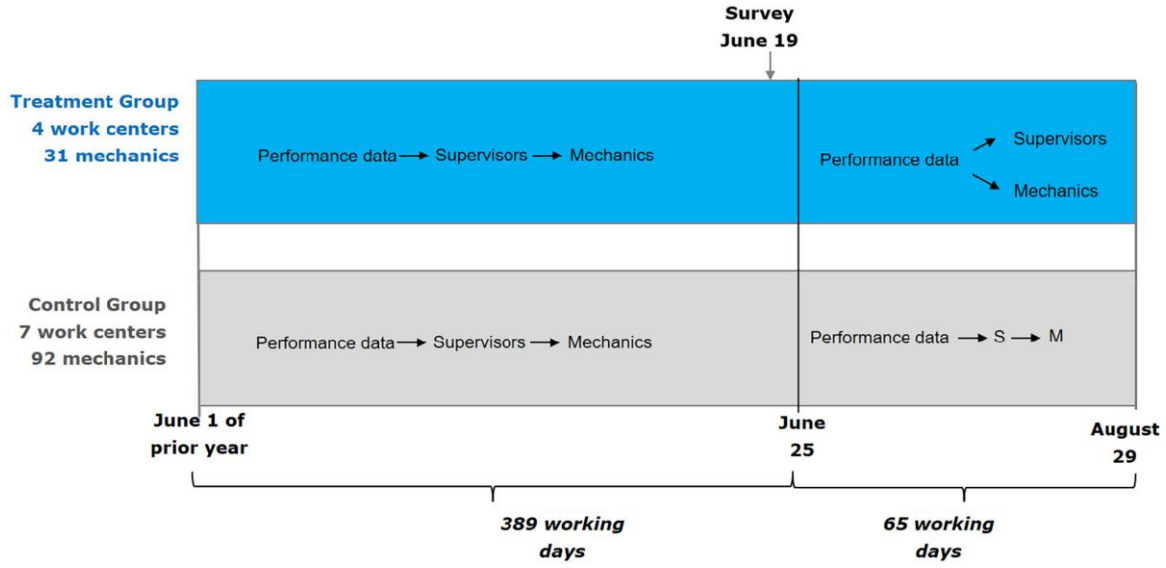


Figure 1. (Color online) Randomization and Timeline



Note. S, supervisors; M, mechanics.

characteristics from those who responded to the survey and those who did not (e.g., tenure, age, race, preintervention average nonproductive/productive time). The difference between respondents and nonrespondents is not statistically significant on any of these measures. This alleviates concerns about nonresponse bias and the generalizability of the findings. Table 1 shows the definitions and descriptive statistics for our key variables, and Table 2 tabulates the unconditional correlations between them.

3.2.2. Dependent Variables. GasCo stringently tracked individuals' time use via three standardized performance metrics: % *productive time*, % *support time*, and % *nonproductive time* (Figure 2). The mechanics each had sole accountability for their own time allocation across these categories, which collectively represent the entire workday. GasCo designed the metrics to help mechanics allocate more time to productive activities. Our decision to use these time measures as performance measures was therefore primarily driven by GasCo's own focus, which dovetailed with our interests.

Productive time was time spent either en route to a customer job or onsite conducting the work. It was heavily constrained by the system's data, logic, and machine-learning algorithms. For example, productive en route time was calculated in light of real-time traffic data. That is, a trip from here to there counted as 20 minutes of productive time if that was how long the traffic data indicated such a trip should take—even if it actually took more than 20 minutes.[6] Support time—maintenance (such as vehicle or building upkeep), training (such as safety briefings), preparation (such as picking up materials and loading or unloading a truck), and colleague support (such as meetings or peer training)—was similarly constrained.

Although mechanics generally understood how these metrics were calculated, they did not understand the algorithm enough to know the exact time an activity "should" take in a particular situation. Neither did they manually tag their time to one of the three categories in the system; their daily work time was automatically tracked and coded. To further prevent abuse or gaming (see Table B.1 for qualitative evidence of

Table 1. Dependent and Independent Variables

Variable	Definition	Mean	Standard deviation	Median
% *Nonproductive Time*	A mechanic's nonproductive hours as a percentage of the total available working hours.	30.696	23.406	23.471
% *Productive Time*	A mechanic's productive hours as a percentage of the total available working hours.	59.125	24.360	66.112
% *Support Time*	A mechanic's support hours as a percentage of the total available working hours.	10.179	16.203	4.624
Tenure	Number of years the mechanic had worked in the company at the start of the intervention.	18.521	8.281	19.000
Age	The mechanic's age (in years) at the beginning of the intervention.	47.206	8.183	50.000
White	=1 if the mechanic's ethnic group is White.	0.552	0.497	1.000
Supervisor Support	Sum of the scores for questions on supervisor support (Q22–Q27 in Appendix C).	24.164	4.007	24.000
Social Comparison Orientation	Sum of the scores for questions on social comparison orientation (Q11–Q21 in Appendix C).	35.193	7.653	37.000
Intrinsic Motivation	Sum of the scores for questions on intrinsic motivation (Q6–Q10 in Appendix C).	16.527	2.362	16.000
Extrinsic Motivation	Sum of the scores for questions on extrinsic motivation (Q1–Q5 in Appendix C).	14.809	2.381	15.000
Prior Performance (Nonproductive)	The average % *Nonproductive Time* in the pre-intervention period.	30.643	12.555	28.642
Prior Performance (Productive)	The average % *Productive Time* in the pre-intervention period.	59.554	12.392	62.942
Prior Performance (Support)	The average % *Support Time* in the pre-intervention period.	9.802	5.858	9.022
Self-evaluation	The self-reported percentile of the pre-intervention performance on % *Productive Time* (Q28 in the Appendix C).	67.331	31.705	81.890

Note. $n = 11{,}120$.

Table 2. Correlations

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. % Nonproductive Time	1.000													
2. % Productive Time	**−0.771**	1.000												
3. % Support Time	**−0.286**	**−0.0390**	1.000											
4. Tenure	−0.003	**−0.042**	**0.068**	1.000										
5. Age	**0.040**	**−0.029**	**−0.015**	**0.706**	1.000									
6. White	**0.074**	**−0.138**	**0.100**	**0.031**	**−0.297**	1.000								
7. Supervisor Support	**−0.052**	**0.035**	0.022	**−0.187**	**−0.222**	**0.113**	1.000							
8. Social Comparison Orientation	**0.026**	**0.029**	**−0.082**	**−0.180**	**−0.049**	**−0.059**	**0.482**	1.000						
9. Intrinsic Motivation	−0.003	**−0.072**	**0.113**	**0.037**	**−0.124**	**−0.055**	**0.410**	**0.305**	1.000					
10. Extrinsic Motivation	0.024	**−0.036**	0.020	**−0.113**	**−0.067**	**−0.139**	**0.401**	**0.580**	**0.593**	1.000				
11. Prior Performance (Nonproductive)	**0.537**	**−0.447**	**−0.104**	0.008	**0.102**	**0.133**	**−0.102**	**0.069**	0.003	**0.041**	1.000			
12. Prior Performance (Productive)	**−0.477**	**0.508**	**−0.075**	**−0.079**	**−0.070**	**−0.256**	**0.094**	**0.042**	**−0.141**	**−0.065**	**−0.890**	1.000		
13. Prior Performance (Support)	**−0.143**	**−0.116**	**0.380**	**0.152**	**−0.069**	**0.255**	0.020	**−0.236**	**0.291**	**0.051**	**−0.261**	**−0.208**	1.000	
14. Self-evaluation	**−0.117**	**0.054**	**0.088**	**−0.191**	**−0.088**	**−0.160**	**0.104**	0.022	**0.165**	**0.099**	**−0.204**	**0.115**	**0.193**	1.000

Notes. $N = 11{,}120$. All correlations in bold are significant at $p < 0.01$.

the low likelihood of gaming), time allocation and activity records were audited. Indeed, one reason GasCo thought more transparent data might be beneficial for the company was that it might reduce the possibility of future abuse by increasing the number of eyeballs on the data.

All three GasCo performance metrics—% productive time, % support time, and % nonproductive time—were in use long before our study, as was the automated system that tracked them; the study was solely focused on the effects of making the resulting performance data transparent to mechanics. In our empirical analysis, % *nonproductive time*, % *productive time*, and % support time are the dependent variables. Throughout the period of our study, there was enough demand for both productive and support activities to keep all mechanics busy 100% of the time with either, leaving them with the option to improve any of their metrics and flexibility in how to do so. From Table 1, we see that throughout the data analysis period, the average % *nonproductive time* is 30.7%, the average % *productive time* is 59.1%, and the average % *support time* is 10.2%. Standard deviation values suggest considerable variation in these dependent variables.

3.2.3. Moderating and Control Variables. To examine moderating variables, we administered a pre-experimental survey (Appendix C) to all the mechanics in our sample.

To measure perceived supervisor support, like Eisenberger et al. (2002) and Shanock and Eisenberger (2006), we used a six-item instrument (Q22–Q27 in Appendix C). Consistent with that instrument's design and previous use, our metric is the sum of the scores from each relevant survey question, adjusted for reverse-coding. Cronbach's alpha for these survey questions is 0.9254.

To measure social comparison orientation, we asked the mechanics to complete the Iowa-Netherlands Comparison Orientation Measure (Gibbons and Buunk 1999, Buunk and Gibbons 2007). Consistent with that well-established scale's design and previous use, our metric for social comparison orientation is the sum of the scores from each relevant survey question (Q11–Q21 in Appendix C), adjusted for reverse coding. Cronbach's alpha for these questions is 0.8824.

To measure intrinsic and extrinsic motivation, in accordance with prior scholarship (for a review, see Mayer et al. (2007)), we used the standard five-item instrument for intrinsic motivation and five-item instrument for extrinsic motivation from the Work Preference Inventory (WPI), a well-established and highly cited instrument for measuring motivation at work (Amabile et al. 1994, Fischer et al. 2019). Although we have no reason to expect that intrinsic and extrinsic motivation will operate orthogonally in our context (as opposed to like two ends of a spectrum), we measure and analyze them separately to allow for that possibility. We think it is important to do so given that the original WPI results showed them to be largely orthogonal (Amabile et al. 1994, p. 958) and that many studies since—including those drawing on the WPI (Choi 2004, Prabhu et al. 2008, Abuhamdeh and Csikszentmihalyi 2009, Robinson et al. 2014), those using other integrated scales (Guay et al. (2000), Situational Motivational Scale; Tremblay et al. (2009) (Work Extrinsic and Intrinsic Motivational Scale); Gagne et al. (2010) (Motivation at Work Scale)), and those using entirely distinct instruments (Ashkanani et al. 2022, Roberts et al. 2006)—have largely

measured and analyzed intrinsic and extrinsic motivation separately on the view that they may be orthogonal (Cameron and Pierce 1994, 1996; Lepper and Henderlong 2000) or even positively correlated (Hennessey et al. 1989, Hennessey and Zbikowski 1993) in some contexts. Our metric for intrinsic motivation is the sum of the scores from the five survey questions related to it (Q6–Q10 in Appendix C), whereas the metric for extrinsic motivation is the sum of the scores from the five questions related to it (Q1–Q5 in Appendix C), adjusted for reverse coding. Cronbach's alpha is 0.7723 for the questions related to intrinsic motivation and 0.6352 for those related to extrinsic motivation.

We control for past performance in all regressions, because people with different "starting points" in performance are likely to respond differently to any intervention designed to improve it. For example, feedback can motivate poorer performers while having little influence on better ones (Pritchard et al. 1981). Because some past-performance effects could be triggered either by actual past performance or by self-assessed past performance (Northcraft and Ashford 1990), we incorporate both as controls.[7] Our measure of a mechanic's actual past performance comes from archival data (averaged over the preintervention period) on % *nonproductive time*, % *productive time*, and % *support time*. Our measure of self-assessed past performance comes from a survey question (Q28 in Appendix C). We standardized the responses on a scale of 0 to 100. By pairing a mechanic's response with actual preintervention performance, we could control for past performance, both actual and perceived (self-assessed) (Meyer 1995). We also incorporate demographic control variables—tenure, age, and race—based on GasCo's internal records (on average, mechanics had 18.5 years of tenure and were 47.2 years old, and ~55% were White).

From Table 1, we see that % *productive time* is negatively correlated with tenure and age and, conversely, % *nonproductive time* is positively correlated with tenure and age (suggesting a decrease in productive efforts with the increase of seniority). We also see that % *productive time* is positively correlated with perceived supervisor support and self-evaluation. We control for all these variables or individual fixed effects in our regressions. Although we do not explicitly control for work quality, GasCo did monitor customer ratings, and the degree to which a customer's issues were addressed on the first visit. Neither of these customer metrics varied significantly during the period of our study.

3.2.4. Participant Observation. To gather qualitative evidence and understand how mechanics and supervisors felt about the intervention, we arranged for one research assistant to be embedded into the workforce for the second week of the experiment. He was selected for his ability to fit in with recruits for the mechanic role at GasCo but was trained to collect field notes (Emerson et al. 1995) and in the fundamentals of participant-observation. For a week, he rode, worked, ate, and hung out with other mechanics as a typical apprentice, rotating among work centers. His notetaking was not seen as unusual but rather as typical for an apprentice. Because mechanics spend so much time driving, there is a lot of time for casual conversation; the newly disclosed transparent performance data were a natural and frequent topic in the treatment work centers. This qualitative evidence, coded by the research question and moderators examined, adds significant texture to the quantitative results of the field experiment and the survey. In particular, it provides insight into why employees' performance changed as revealed by the quantitative analyses.

3.3. Main Regression Model

We visualize the impact of the intervention in Figure 3. Up to the intervention, the weekly moving average lines showed approximately parallel trends between treatment and control groups for the three performance variables. To formally analyze the field experiment data, we used a difference-in-differences model (Meyer 1995) to estimate the effect of an intervention on treatment units relative to control units during the same period. Our model is

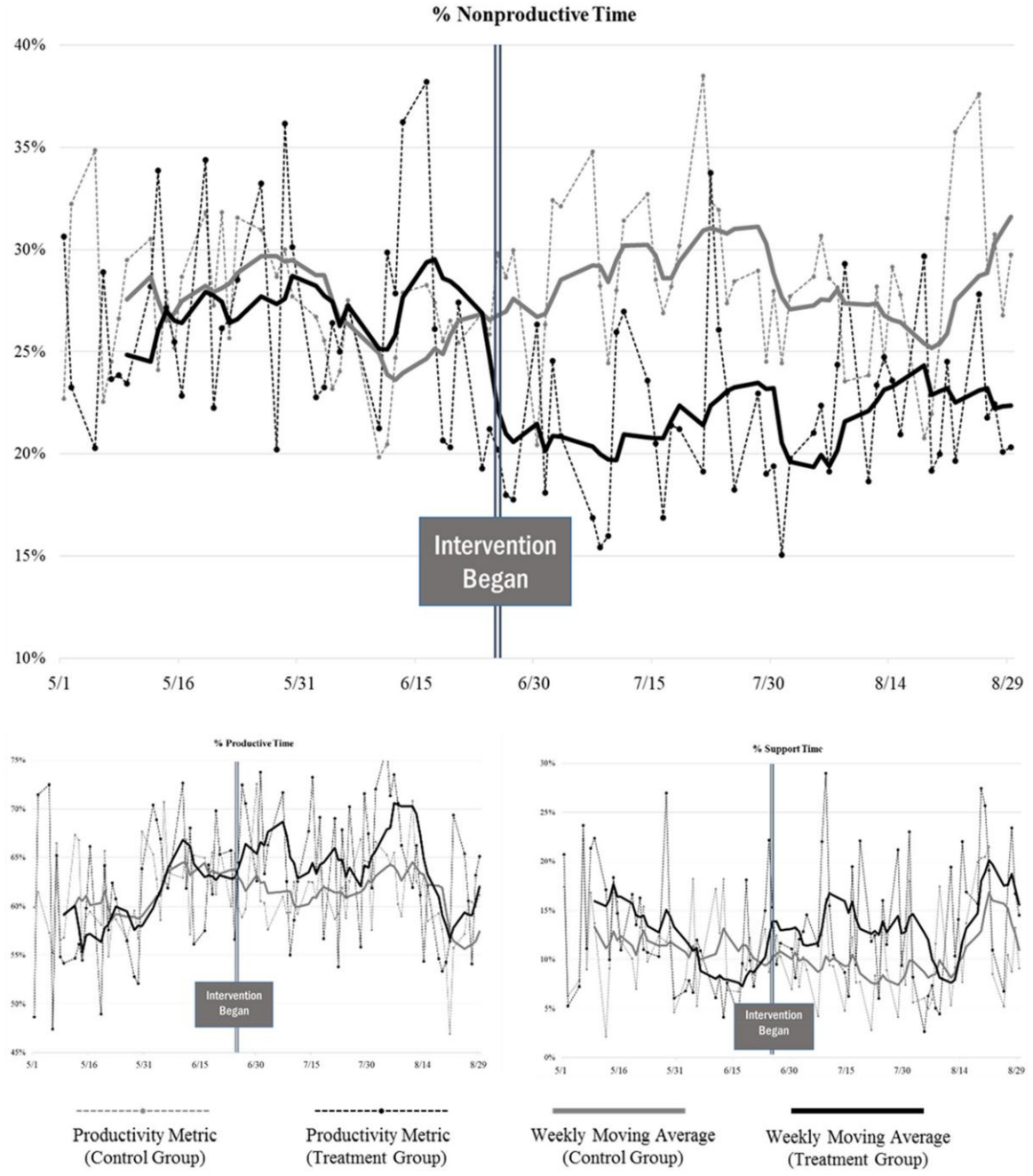
$$Y_{it} = \alpha + (\beta_1 \times Treatment_{it}) + (\beta_2 \times Post_{it}) + (\beta_3 \times (Treatment_{it} \times Post_{it})) + \sum Controls + \varepsilon_{it}$$

or

$$Y_{it} = \alpha + (\beta_2 \times Post_{it}) + (\beta_3 \times (Treatment_{it} \times Post_{it})) + Individual\ Fixed\ Effects + \varepsilon_{it},$$

where Y_{it} is the performance metric at the employee(i)-workday(t) level, $Treatment_{it}$ is an indicator variable that equals one if the employee worked in one of the work centers in the treatment group, and $Post_{it}$ is an indicator variable that equals one if the date was on or after June 25, when the intervention began. The main estimation uses ordinary least squares (OLS) regressions. Consistent with Bertrand et al. (2004), standard errors of the coefficients are corrected for autocorrelation and clustered by work center.[8] If the treatment effect is positive, we should see a negative and significant β_3 when Y_{it} is the negative productivity indicator, % *nonproductive time*, and a positive and significant β_3 when Y_{it} is the positive productivity indicator, % *productive time*. In testing the average

Figure 3. Visualization of Treatment Effects



treatment effect, we ran two specifications: one including as *Controls* all the available employee-level control measures (*Tenure*, *Age*, *Supervisor Support*, etc.) and the other including individual fixed effects (absorbing the treatment indicator and all our control measures and controlling for unobservable time-invariant individual characteristics). In subsequent analyses, we use only the stricter specification (individual fixed effects).

4. Results

4.1. How Does Partial Disintermediation of Supervisors via Transparent Performance Data Affect Front-Line Performance?

Table 3 shows the main regression results with and without individual fixed effects. Columns 1 and 4 use % *nonproductive time* as the dependent variable, columns 2 and 5 use % *productive time*, and columns 3 and 6 use % *support time*. The coefficient on *Treat* × *Post* is

Table 3. Does Access to Transparent Performance Data Improve Performance?

	(1) % nonproductive time	(2) % productive time	(3) % support time	(4) % nonproductive time	(5) % productive time	(6) % support time
Treat × Post	−3.410*	2.846	0.576	−3.397*	2.697	0.700
	(0.013)	(0.412)	(0.826)	(0.010)	(0.438)	(0.792)
Treat	0.299	−0.081	−0.146			
	(0.232)	(0.728)	(0.409)			
Post	1.264	−3.528***	2.267^	1.349	−3.589***	2.240^
	(0.250)	(0.000)	(0.096)	(0.229)	(0.000)	(0.089)
Tenure	0.023	−0.048*	0.022			
	(0.214)	(0.033)	(0.206)			
Age	−0.061*	0.037	0.022*			
	(0.019)	(0.187)	(0.037)			
White	−0.326	−0.071	0.218			
	(0.404)	(0.865)	(0.230)			
Supervisor Support	0.076	−0.148^	0.076*			
	(0.361)	(0.087)	(0.010)			
Social Comparison	−0.072^	0.070*	0.013			
	(0.069)	(0.011)	(0.633)			
Intrinsic Motivation	−0.165*	0.152^	−0.028			
	(0.047)	(0.058)	(0.717)			
Extrinsic Motivation	0.182^	−0.135	−0.054			
	(0.054)	(0.119)	(0.699)			
Prior Performance (Nonproductive)	1.005***					
	(0.000)					
Prior Performance (Productive)		0.999***				
		(0.000)				
Prior Performance (Support)			1.044***			
			(0.000)			
Self-evaluation	−0.007^	−0.004	0.010*			
	(0.086)	(0.343)	(0.023)			
Individual fixed effects?	No	No	No	Yes	Yes	Yes
Observations	11,120	11,120	11,120	11,120	11,120	11,120
Adjust R^2	0.29	0.26	0.15	0.29	0.26	0.15

Notes. This table reports OLS regression results of % *nonproductive time* (columns 1 and 4), % *productive time* (columns 2 and 5), and % *support time* (columns 3 and 6) on a treatment indicator (*Treat*), a postintervention indicator (*Post*), an interaction of the two variables (*Treat* × *Post*), and other controls. In columns 4, 5 and 6, *Treat* is absorbed by individual fixed effects and is not reported. All variables are defined in Table 1. Standard errors are clustered at the work-center level, and *p*-values are reported in parentheses.

^$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

the estimated treatment effect of the intervention on the performance metric. The negative coefficients on this interaction term in columns 1 ($\beta = -3.41$, $p = 0.013$) and 4 ($\beta = -3.397$, $p = 0.010$) indicate a negative treatment effect on % *nonproductive time*; that is, mechanics in the treatment group spent less nonproductive time after the intervention than those in the control group did. Compared with the preintervention mean value, the results in column 4 indicate an 11% decrease in % *nonproductive time*, suggesting that access to transparent performance data previously available only to supervisors improved performance. Mechanics felt they got "the same information as before" (M5) but generally agreed that "this is better. I get to see the data now, not when [supervisor] feels like it. I get all of it, not just what [supervisor] remembers. I get the numbers, not just [supervisor]'s words" (M2).[9] Table B.1 provides more qualitative evidence on how mechanics and supervisors reacted to the intervention.

Columns 2 and 5 of Table 3, however, present an important caveat to the improved performance. Comparing columns 1 and 4 with 2 and 5, the treatment effect on nonproductive time is greater than that on productive time, in both magnitude and statistical significance; mechanics focused on reducing nonproductive time and not on increasing productive time. Figure 3 shows that after the intervention, the treatment group showed a significant decrease in nonproductive time relative to the control group, larger than the relative increase in productive time.[10] This suggests that access to transparent performance data triggers, on average, a greater behavioral shift toward avoiding nonproductive behaviors than toward approaching strictly productive behaviors. As perceived by the mechanics and captured

by the participant-observer, it made almost all the mechanics more determined *not* to stand out for bad, or nonproductive, activities, but not more determined to stand out for good, or productive, activities. In their own words, the goal was to "hide in the middle of the pack" (M4) and to "conform, not excel" (M16). As one explained, "No one hears anything about being middle of the good [productive] time … but if your bad [nonproductive] time is high, I think people notice … Hell, I notice when others are high …. When it was just [supervisor name], it was different—he got me, he knows what it's like, that the good is more important than the bad—but the numbers don't, so just be sure the bad doesn't stand out" (M3). Table B.1 provides more qualitative evidence. Other research has shown that motivation to conform and comply differs from motivation to excel (Cialdini and Trost 1998), and in our study, the substitution of transparent performance data for supervisor-intermediated traditional performance feedback seemed to favor the former source of motivation for the average worker.

To check the quality of our intervention—that is, to confirm that access to transparent performance data and not something unrelated drove the results—we obtained the number of views of the scorecards for each treatment site. Across the four sites, the average employee accessed the report at least once every three days and some far more frequently. Sites whose mechanics accessed the report more often had the greater improvement (in reducing nonproductive time *and* increasing productive time), suggesting that it was, indeed, transparent performance data that drove the effects (see Appendix D for the additional analyses).

4.2. Moderating Role of Supervisor Quality (Vertical Relational Moderator)

To investigate the role of supervisor quality, we measured it with perceived supervisor support (from the pre-experimental survey) and ran the baseline regressions, following the precedent set by prior studies by splitting the sample (Baron and Kenny 1986, Jaccard and Turrisi 2003) into those who reported perceived supervisor support that was (a) greater than or equal to the sample median and those who reported it (b) below the median; these groups are labeled "High Supervisor Support" and "Low Supervisor Support" in Table 4. This analysis examines whether it was mechanics who saw their supervisors as supportive or those who saw them as unsupportive who showed more performance improvement after receiving access to transparent performance data. We find treatment effects of greater magnitude and statistical significance for both % *nonproductive time* and % *productive time* for those who perceived low supervisor support, suggesting that access to transparent performance data could serve as a substitute for "bad" supervisors. The treatment effects on % *support time* are statistically insignificant at the 5% level in both subsamples ($\beta = 1.468$, $p = 0.631$ and $\beta = -4.080$, $p = 0.063$). Compared with the preintervention mean values, we see a 16.27% decrease in % *nonproductive time* ($\beta = -5.298$, $p = 0.023$) and a 16.07% increase in % *productive time* ($\beta = 9.377$, $p < 0.001$) for the low-supervisor-support subsample. Calculating the z-statistics that compare the coefficients on *Treat* × *Post* between columns 3 and 4 (Stock and Watson 2003, Dunn et al. 2012), the difference in treatment effects for % *productive time* between the high- and low-supervisor-support subsamples is

Table 4. Moderating Role of Supervisor Support

	(1) % nonproductive time - high supervisor support		(2) % nonproductive time - low supervisor support		(3) % productive time - high supervisor support		(4) % productive time - low supervisor support		(5) % support time - high supervisor support		(6) % support time - low supervisor support
Treat × *Post*	−2.258	$z = -1.34$	−5.298*		0.790	$z = 2.01$*	9.377***		1.468	$z = 1.59$*	−4.080^
	(0.126)		(0.023)		(0.850)		(0.000)		(0.631)		(0.063)
Post	−0.278		3.004*		−3.537*		−3.670**		3.258*		0.666
	(0.832)		(0.038)		(0.038)		(0.004)		(0.024)		(0.725)
Individual fixed effects?	Yes		Yes		Yes		Yes		Yes		Yes
Observations	7,159		3,961		7,159		3,961		7,159		3,961
Adjusted R^2	0.24		0.36		0.24		0.31		0.24		0.31

Notes. This table reports OLS regression results of % *nonproductive time* (columns 1 and 2), % *productive time* (columns 3 and 4), and % *support time* (columns 5 and 6) on a postintervention indicator (*Post*), an interaction of the two variables (*Treat* × *Post*), and individual fixed effects. "High (Low) Supervisor Support" is the subsample of mechanics who reported a high (low) level of perceived supervisor support—above or equal to (below) the sample median—in the pre-experimental survey. All variables are defined in Table 2. Standard errors are clustered at the workcenter level and p-values are reported in parentheses. Comparing the coefficients on *Treat* × *Post* between columns 1 and 2 yields a z-statistic of −1.34; between columns 3 and 4, a z-statistic of 2.01; and between columns 5 and 6, a z-statistics of 1.59.

^$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

statistically significant at the 5% level. Our participant-observer heard one mechanic say that he liked the fact that "low-performers can approach high-performers" (M10)—knowing now who they were—to learn how to improve if they "didn't have a good supervisor" (M10). Employees with less supervisor support improved more "because mechanics can support each other directly" (M8), whereas, for those employees, discussions with their supervisors had been just "check-the-box exercises" (M6) and "a total waste of time" (M6). Table B.1 provides more qualitative evidence for how mechanics and supervisors reacted to the intervention. Getting past these bad intermediaries is like removing a blood clot—information flows more freely, communication cost is lower, and employees get the signals they need to improve. Our participant-observer heard some employees refer to supervisors who "everyone knows played favorites" (PO), whereas transparent performance data were described as "just about the work" and a "reality check that I can trust" (PO). Although supervisors overall received strong ratings, mechanics who worked for the less-supportive ones benefited more when those supervisors were partially disintermediated by transparent performance data.

4.3. Moderating Role of Social Comparison Orientation (Horizontal Relational Moderator)

We also ran the baseline regressions on social comparison orientation, splitting the sample into those who reported it (a) greater than or equal to the sample median and those who reported it (b) below the median; these groups are labeled "High Social Comparison" and "Low Social Comparison" in Table 5. This analysis examines whether mechanics who have higher social comparison orientation (and hence are more influenced by relative performance information) showed more performance improvement after receiving access to performance data. The treatment effect on % *nonproductive time* is larger in magnitude for the subsample with *low* social comparison orientation ($\beta = -5.526$, $p = 0.005$), and the difference in the treatment effects between the two subsamples (columns 1 and 2) is statistically insignificant, *inconsistent* with relative performance information being the driving force behind this main effect. The treatment effect on % *productive time* is larger in magnitude for the high-social-comparison-orientation subsample, but the treatment effects on both subsamples are statistically insignificant ($\beta = 4.829$, $p = 0.183$ and $\beta = 2.506$, $p = 0.645$, respectively), and the difference between the two subsamples (columns 3 and 4) is statistically insignificant. Treatment effects on support time are insignificant in both subsamples ($\beta = -2.280$, $p = 0.414$ and $\beta = 3.020$, $p = 0.503$, respectively) and not significantly different across the subsamples. We therefore cannot assign social comparison orientation a significant moderating role in the effect of transparent data on performance—that is, it is not a main mechanism through which transparent data affected performance.

There are several possible explanations for these mixed and inconclusive results related to social comparison. In real work environments (as opposed to a laboratory setting), it is quite possible for employees to have a sense of how they compare with each other—from their supervisors' comments and just from observing each other—without any mention of it in formal feedback. To the extent that a person with a greater social comparison orientation would more frequently seek out such informal information already, the effect of providing it

Table 5. Moderating Role of Social Comparison Orientation

	(1) % nonproductive time - high social comparison	(2) % non productive time - low social comparison	(3) % productive time - high social comparison	(4) % productive time - low social comparison	(5) % support time - high social comparison	(6) % support time - low social comparison
Treat × *Post*	−2.549 $z = -1.28$	−5.526**	4.829 $z = 0.37$	2.506	−2.280 $z = -1.05$	3.020
	(0.204)	(0.005)	(0.183)	(0.645)	(0.414)	(0.503)
Post	−0.302	3.970*	−2.446*	−5.403*	2.748	1.434
	(0.814)	(0.023)	(0.041)	(0.021)	(0.136)	(0.168)
Individual fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,909	5,211	5,909	5,211	5,909	5,211
Adjusted R^2	0.29	0.29	0.26	0.26	0.10	0.19

Notes. This table reports OLS regression results of % *nonproductive time* (columns 1 and 2), % *productive time* (columns 3 and 4), and % *support time* (columns 5 and 6) on a postintervention indicator (*Post*), an interaction of the two variables (*Treat* × *Post*), and individual fixed effects. "High (Low) Social Comparison" is the subsample of mechanics who reported a high (low) level of social comparison orientation—above or equal to (below) the sample median—in the pre-experimental survey. All variables are defined in Table 2. Standard errors are clustered at the work-center level and *p*-values are reported in parentheses. Comparing the coefficients on *Treat* × *Post* between columns 1 and 2 yields a *z*-statistic of −1.28; between columns 3 and 4, a *z*-statistic of 0.37; and between columns 5 and 6, a *z*-statistics of –1.05.

^$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

formally could be smaller. Another possibility is that heightened social comparison could lead to negative psychological consequences such as diminished trust in coworkers (Dunn et al. 2012), discouragement (Beshears et al. 2015), and unproductive behaviors. In any case, the results in Table 5—and the qualitative data (Table B.1) that implies, at most, unsystematic interactions between mechanics to collaborate on improvements but little more—suggests that the main effect of our intervention is unlikely to be primarily driven by greater access to relative performance information.

4.4. Moderating Role of Work Motivation Type (Relational vs. Self-Moderator)

We also ran the baseline regressions for intrinsic motivation, splitting the sample into those who reported intrinsic motivation that was (a) greater than or equal to the sample median and those who reported it (b) below the median; these groups are labeled "High Intrinsic Motivation" and "Low Intrinsic Motivation" in Table 6. This analysis examines whether mechanics with lower intrinsic motivation for work showed more performance improvement after receiving access to transparent performance data.

We find treatment effects of greater magnitude and statistical significance for both % *nonproductive time* and % *productive time* for those who had low intrinsic motivation, suggesting that access to transparent performance data could serve as a substitute for intrinsic motivation. As one supervisor observed, "Even though I would talk to [mechanics] about it when there was something concerning in my daily reports, now that they can access it directly, [those mechanics] seem to take it more seriously." (S6) Compared with the preintervention mean values, we see a 21.28% decrease in % *Nnonproductive time* ($\beta = -6.711$, $p < 0.001$) and a 13.85% increase in % *productive time* ($\beta = 8.303$, $p = 0.003$) for the low-intrinsic-motivation subsample. Treatment effects on % *support time* are statistically insignificant in both subsamples ($\beta = 2.652$, $p = 0.418$ and $\beta = -1.592$, $p = 0.417$, respectively). Using z-statistics to compare coefficients on *Treat* × *Post* between columns 1 and 2 (Stock and Watson 2003), the difference in treatment effects for % *nonproductive time* between the high- and low-intrinsic-motivation subsamples is statistically significant at the 1% level. The difference in treatment effects for % *productive time* between these two subsamples is statistically significant at the 5% level. We emphasize that these differences stem from the improvements in the subsample with low intrinsic motivation (columns 2 and 4) and not from a deterioration of performance in the subsample with high intrinsic motivation (columns 1 and 3).

We ran the baseline regressions for extrinsic motivation, as well, splitting the sample into those who reported extrinsic motivation that was (a) greater than or equal to the sample median and those who reported it (b) below the median; these groups are labeled "High Extrinsic Motivation" and "Low Extrinsic Motivation" in Table 7. This analysis examines whether mechanics with higher extrinsic work motivation showed more performance improvement after receiving access to transparent performance data. We find larger treatment effects for both % *nonproductive time* and % *productive time* for those who had high extrinsic motivation. Compared with the preintervention mean values, we see a 15.20% decrease in % *nonproductive time* ($\beta = -4.579$, $p = 0.011$) and a 13.32% increase in % *productive time*

Table 6. Moderating Role of Intrinsic Motivation

	(1) % nonproductive time - high intrinsic motivation		(2) % nonproductive time - low intrinsic motivation		(3) % productive time - high intrinsic motivation		(4) % productive time - low intrinsic motivation		(5) % support time - high intrinsic motivation		(6) % support time - low intrinsic motivation
Treat × *Post*	−0.649	$z = -3.13$**	−6.711***		−2.003	$z = 2.14$*	8.303**		2.652	$z = 1.17$	−1.592
	(0.710)		(0.000)		(0.654)		(0.003)		(0.418)		(0.417)
Post	0.624		2.258^		−3.287*		−3.968**		2.663^		1.710
	(0.666)		(0.061)		(0.013)		(0.002)		(0.061)		(0.280)
Individual fixed effects?	Yes		Yes		Yes		Yes		Yes		Yes
Observations	6,050		5,070		6,050		5,070		6,050		5,070
Adjusted R^2	0.25		0.34		0.23		0.31		0.15		0.15

Notes. This table reports OLS regression results of % *nonproductive time* (columns 1 and 2), % *productive time* (columns 3 and 4), and % *support time* (columns 5 and 6) on a postintervention indicator (*Post*), an interaction of the two variables (*Treat* × *Post*), and individual fixed effects. "High (Low) Intrinsic Motivation" is the subsample of mechanics who reported a high (low) level of intrinsic motivation—above or equal to (below) the sample median—in the pre-experimental survey. All variables are defined in Table 2. Standard errors are clustered at the work-center level, and p-values are reported in parentheses. Comparing the coefficients on *Treat* × *Post* between columns 1 and 2 yields a z-statistic of −3.13; between columns 3 and 4, a z-statistic of 2.14; and between columns 5 and 6, a z-statistic of 1.17.

^$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 7. Moderating Role of Extrinsic Motivation

	(1) % nonproductive time - high extrinsic motivation	(2) % nonproductive time - low extrinsic motivation	(3) % productive time - high extrinsic motivation	(4) % productive time - low extrinsic motivation	(5) % support time - high extrinsic motivation	(6) % support time - low extrinsic motivation
Treat × *Post*	−4.579* $z = -1.01$	−2.858*	7.875*** $z = 1.75$^	−0.754	−3.296^ $z = -1.49$	3.612
	(0.011)	(0.011)	(0.000)	(0.878)	(0.082)	(0.425)
Post	1.134	1.682^	−3.822**	−3.227**	2.688	1.545
	(0.481)	(0.063)	(0.003)	(0.004)	(0.115)	(0/158)
Individual fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,054	5,066	6,054	5,066	6,054	5,066
Adjusted R^2	0.25	0.34	0.26	0.36	0.15	0.15

Notes. This table reports OLS regression results of % *nonproductive time* (columns 1 and 2), % *productive time* (columns 3 and 4), and % *support time* (columns 5 and 6) on a postintervention indicator (*Post*), an interaction of the two variables (*Treat* × *Post*), and individual fixed effects. "High (Low) Extrinsic Motivation" is the subsample of mechanics who reported a high (low) level of extrinsic motivation—above or equal to (below) the sample median—in the pre-experimental survey. All variables are defined in Table 2. Standard errors are clustered at the work-center level, and *p*-values are reported in parentheses. Comparing the coefficients on *Treat* × *Post* between columns 1 and 2 yields a *z*-statistic of −1.01; between columns 3 and 4, a *z*-statistic of 1.75; and between columns 5 and 6, a *z*-statistic of –1.49.

^$p < 0.1$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

($\beta = 7.895$, $p < 0.001$) for the high-extrinsic-motivation subsample. Treatment effects on % *support time* are statistically insignificant at the 5% level in both subsamples ($\beta = -3.296$, $p = 0.082$ and $\beta = 3.612$, $p = 0.425$, respectively). Calculating the *z*-statistics that compare the coefficients on *Treat* × *Post* between columns 3 and 4 (Stock and Watson 2003), the difference in treatment effects for % *productive time* between the high- and low-extrinsic-motivation subsamples is statistically significant at the 10% level. However, the difference in treatment effects for % *nonproductive time* between these two subsamples is statistically insignificant. This is unsurprising given that both subsamples significantly reduced nonproductive time (columns 1 and 2), whereas only the subsample with higher extrinsic motivation (column 3) significantly increased productive time.

4.5. Robustness Checks

We ran robustness checks to address four characteristics of our field experiment. First, the gas utility business is seasonal, which can be seen, in some regressions, in the significant coefficient on our time variable, *Post*. In part to account for seasonality, we requested a much longer time series of preintervention performance data and reran the regressions using month fixed effects. The results were similar. Second, because our dependent variables—% *nonproductive time*, % *productive time*, and % *support time*—are correlated, the error terms in the two regressions are correlated. Correlated dependent variables do not cause bias in the estimation of coefficients, but running them as separate regressions could reduce efficiency (Kennedy 2003). Therefore, we also used seemingly unrelated regression estimation (SURE) to estimate our main regressions; results did not change. Third, our dependent variables have bounded values. We ran Tobit regressions, setting the lower and upper bounds at 0 and 100, and saw similar results. Fourth, a mechanic's daily performance data are correlated with past performance. In our reported analysis, we controlled for preintervention performance (actual and self-assessed) and clustered standard errors by work location (to account for the correlation among all mechanics' performance data within the same work location, as employees in the same location receive the same treatment condition and may influence each other's performance). As a robustness check, we instead ran the regressions clustering standard errors by individual and without the self-assessed performance variable. In both cases, results were similar.

5. Conclusion

Our goal in this study was to examine how providing employees with direct digital access to performance data previously available only to supervisors affects front-line results. In a field experiment in a large U.S. service organization, our findings indicate that, on average, partially disintermediating supervisors through data transparency does improve employee performance. Yet there are two important nuances to note. First, access to the data did not encourage an overall increase in the "best" (productive) behaviors as much as it encouraged avoidance of the "worst" (unproductive) ones—In short, it triggered, on average, more avoidance than approach

behaviors (Roth and Cohen 1986, Elliot and Church 1997, Elliot and Thrash 2002). Second, access produced significant performance improvements primarily for certain employee subsamples—those who perceived their supervisors as less supportive, those with low intrinsic motivation, and those with high extrinsic motivation—with implications for how we conceptualize technology-delivered versus supervisor-intermediated performance feedback. Our study highlights the impact of those relational moderators on performance, extending the application of FIT (traditionally used to explain the circumstances under which supervisor-intermediated feedback becomes more effective) to settings where supervisors serve as "apomediaries" rather than gatekeepers. In so doing, it also suggests—perhaps—that the channel of data delivery (computer or human) may more generally affect how FIT findings ought to be applied.

We identify a causal relationship between data transparency and employee performance using a field experiment. However, as in any field experiment of this complexity, identifying causal relationships comes at a cost. The single U.S. setting, single-period intervention, single-profession subject pool, and relatively independent work, although enhancing the reliability of our analyses, also restrict the broader generalizability of our findings. Therefore, it is important to understand the boundaries of our results. Future field research can extend our study by exploring how performance effects—and especially the relevant moderators—might change in other organizational cultures, in other professions, in settings with more interdependent work, or in environments where "good," "bad," and "neutral" uses of time are defined differently. Furthermore, laboratory research is well positioned to more generally explore, beyond our field context, whether providing direct access to performance data enhances avoidance motives, without a countervailing effect on approach motives. In addition, future studies can investigate other moderators that might make the direct-access intervention more successful, such as how supervisors are trained to give feedback in the new system and ways they can help the traditional and new systems work well together. Researchers can also identify and examine particular "value-adding" roles that supervisors-as-apomediaries might play in the feedback process.

Harkening back to the time-use studies of scientific management, our main result prompts us to wonder if this emerging managerial approach—giving workers direct access to performance data rather than aggregating, filtering, and framing it first—improves outcomes as much as organizations expect. We also wonder what other types of employees (beyond the front-line employees we studied) might benefit from data transparency versus manager-intermediated feedback. Especially in the wake of the COVID-19 pandemic, when a fast transition to widespread remote work has prompted many organizations to scramble for ways to supervise and manage using transparent performance data rather than face-to-face interactions, our results provide some reason to pause and question how this shift in approach might be affecting differently situated employees differently. Our field experiment shows that performance benefits depend on how supportive the supervisor already is, on the employee's motivation type, and on whether the measured activities are productive (billable), nonproductive (nonbillable), or support (future-investment-related) tasks. The first result suggests one mechanism triggered by partial disintermediation of supervisors: when employees value the data source more than they value the middleman (but not vice versa), performance improves. The second and third results suggest a second mechanism: When employees can manage their motivation type better than intermediating supervisors can filter for it, they can make better decisions about where to spend their time, and performance improves. In short, there is power in access to more transparent performance data without supervisor intermediation, but neither scholars nor practitioners should expect the resulting performance benefits to emerge as desired without taking into account idiosyncratic differences in employees, in supervisor-employee relationships, and in the types of performance data being shared.

Appendix A.

Table A.1. Review of Existing Empirical Studies of Computer-Mediated vs. Human-Delivered Employee Feedback[a]

Author (yr)	Title	Journal	Laboratory vs. field	Subject population	Relevant feedback delivery treatment/ control difference	Feedback description	Feedback content same across mediums?	Dependent variable(s)	Finding(s)
Alder (2007); Alder and Ambrose (2005)	Examining the relationship between feedback and performance in a monitored environment: A clarification and extension of feedback intervention theory; An examination of the effect of computerized performance monitoring feedback on monitoring fairness, performance, and satisfaction	The journal of high technology management research; organizational behavior and human decision processes	Laboratory	Undergraduate students	Medium[b]: (1) computer, (2) supervisor	No. catalog orders entered (quantity), error rate (quality), destructive versus constructive feedback message	Yes	Interpersonal fairness, desire to respond, task performance (quantity and quality), job satisfaction	Interpersonal fairness ratings higher when receiving *constructive* face-to-face feedback from supervisor versus computer, but not when receiving *destructive* feedback
Earley (1988)	Computer-generated performance feedback in the magazine-subscription industry	Organizational behavior and human decision processes	Field	Magazine publisher employees	Medium: (1) supervisor, (2) computer	No. Subscriptions Processed Relative to Goal (Specific vs. General)	Yes[c]	Trust in feedback, task performance	Feedback will be more effective in increasing performance if received from computer vs. Supervisor
Hebert and Vorauer (2003)	Seeing through the screen: is evaluative feedback communicated more effectively in face-to-face or computer-mediated exchanges?	Computers in human behavior	Laboratory	Undergraduate students	Medium: (1) face-to-face, (2) email	Judge's evaluation of essay written by participant	Yes	Skill appraisal, likeability, and personality traits (rated by judge, participant's estimate of judge's ratings, and participant's self-ratings)	Judges communicated more positive feedback on skills when face-to-face than over email, but the effect did not extend to feedback on liking or traits
Kluger and Adler (1993)	Person- vs. computer-mediated feedback	Computers in human behavior	Laboratory	Undergraduate engineering students	Medium: (1) person, (2) computer	Correct/Incorrect	Yes	Motivation (including time spent on optional math questions), performance	Subjects more likely to seek feedback from a computer than from a person
Kurtzberg et al. (2006)	The effect of email on attitudes toward performance feedback	International journal of organizational analysis	Laboratory	MBA students	Medium: (1) email, (2) paper, (3) face-to-face	Feedback about managerial performance containing positive and negative elements	Yes	Level of negativity in response to feedback, perception of feedback as developmental and evaluative	Subjects responded most positively to feedback via paper, then face-to-face, then email

Table A.1. (Continued)

Author (yr)	Title	Journal	Laboratory vs. field	Subject population	Relevant feedback delivery treatment/ control difference	Feedback description	Feedback content same across mediums?	Dependent variable(s)	Finding(s)
Nargesi (2021)	Comparing AI and human feedback: The role of source in feedback perceptions	Ma thesis	Laboratory	MTurk panelists	Source: (1) ai rating system, (2) human rater	Correct/On track/ Incorrect for each response participants selected to a management situational judgment test, plus objective vs. subjective written commentary	Yes	Feedback trust, source credibility, interpersonal justice, feedback quality	Human rater considered more trustable source only when feedback was subjective; no difference when feedback objective or in terms of source credibility, interpersonal justice, and feedback quality
Northcraft and Earley (1989)	Technology, credibility, and feedback use	Organizational behavior and human decision processes	Laboratory	MBA students	Medium: (1) organization, (2) supervisor, (3) self-generated with computer, or (4) self-generated without computer	3 pages of feedback including record of prior transactions' impact on portfolio value, number of correct predictions of each stock, and update on new prices/ recommendations for each stock	Yes	Credibility of feedback (trust, reliability, usefulness, accuracy), task performance	Self-generated feedback (with or without computer) influenced credibility of feedback, strategy acquisition, and performance, vs. Feedback from supervisor
Thuillard et al. (2022)	When humans and computers induce social stress through negative feedback: Effects on performance and subjective state	Computers in human behavior	Laboratory	Nonpsychology students	Source: (1) computer, (2) human, (3) no feedback	Negative task performance feedback on paper printed from computer or handwritten by human	Yes	Performance, affect, interpersonal fairness, desire to improve, self-esteem	Computer feedback rated as being less fair than human feedback

[a]Note on scope of literature review: Although only one of the studies in this table is a true field experiment of supervisors and their employees, this review adopted an expansive scope to incorporate as many empirical studies as possible that might be relevant to our examination of computer-mediated versus human-delivered employee feedback. Incorporated are (a) a number of laboratory experiments from the management and organizational behavior literatures that simulate a supervisor giving an employee feedback and (b) three *Computers in Human Behavior* studies that were not specific to a work environment, but that were structured closely in experimental design and thus could be relevant to our research question (Kluger and Alder 1993). Potentially similar empirical work in more distant literature, such as education scholars examining students receiving school performance feedback or medical scholars examining college students receiving alcohol usage feedback, produced a number of empirical studies on computer-mediated feedback that seemed potentially instructive, but after we looked at those studies more carefully, they all turned out to be different enough that they were too disconnected from the research question of this work to be included. (Examples include papers focused only on computer-mediated feedback relative to no feedback (Neighbors et al. 2004, Neri et al. 2008) without any comparison with human-delivered feedback; papers focused only on understanding what content (Weitzel et al. 2007, Corbalan et al. 2010, Murphy 2010, Johnson et al. 2015) and timing (Van der Kleij et al. 2015) of computer-mediated feedback was most effective, again without any comparison with human-delivered feedback; and papers comparing computer-mediated versus human-delivered feedback interventions where the content of the two was very different (rather than based on the same set of metrics as in our study), such as college students receiving an Alcohol 101 CD-ROM versus participating in a Brief Motivational Interview (Barnett et al. 2007) or health center visitors receiving a web-based assessment with personalized feedback versus a cognitive behavioral therapy session (Donohue et al. 2004).

[b]Feedback source refers to who/what *generated* the feedback, whereas feedback medium refers to who/what *delivered* the feedback (Thuillard et al. 2022).

[c]"In the supervisor source condition, the workers were told that the supervisor used the computer data to analyze and inform the worker as to their performance level. In the self-source condition, the workers analyzed their own performance data from the computer" (Earley 1988, p. 56).

Appendix B.

Table B.1. Illustrative Qualitative Data (from Participant-Observation) Capturing the Reactions from Mechanics and Supervisors in the Treatment Group to Providing Employees with Transparent Performance Data

Category	Subcategory	Illustrative quotes
Context	General responses to intervention	"We still have the same conversations with [supervisor name] as before. Now that we have the reports, we can skip the reveal and go right to planning... but we still talk about it. Nothing has really changed, you know...." (M4)
		"I get the same information as before. No big deal, these reports. It's like choosing between a self-service and full-service gas station (laughs) …. To be honest, I think [my supervisor] probably shared all this [points at the dashboard] with us in our daily meetings. But sure I prefer having direct access to it. I like knowing what he sees." (M5)
		"I know I could get this information before—my supervisor told me I could just ask—but asking takes time and can also make you look bad. There's no reason why I should have to ask. I don't want to wait to get information until it's too late to use it. This new way is much better." (M1)
		"I finally know what management sees about me! Now I understand some of the feedback I've gotten in the past." (M7)
		"Of course this is better. I get to see the data now, not when [supervisor] feels like it. I get all of it, not just what [supervisor] remembers. I get the numbers, not just [supervisor]'s words." (M2)
		"The dashboards—those numbers aren't usually any surprise. We all know how we are doing." (M9)
		"Nothing has changed with the dashboards. They all get the same information as before. [The dashboards] are just a different format." (S2)
Context	Comments on gaming the system (prompted by participant-observer questions about if it's possible)	"The system can't really be gamed. It's all automatically recorded and stuff. Maybe that's why they are doing this dashboard thing–to make it more visible to us?" (M4)
		"[GasCo] worked with us [the union] to design a system that could not be gamed—by us or by them.... Everything is automatic." (M16 / a union leader)
		"The guys know the numbers are the numbers. I mean, corporate did a solid job at making sure the system is not gameable. I was at a company before where you could play the system, but not here..." (S3)
		"You know, you can't play with the numbers... bad numbers, bad day; good day, good numbers." (S4)
Moderators	Supervisor quality/support	"I still ask [my supervisor] about the dashboards, so I still learn from her advice. But this is better–it is good for us to click in and see it like she does." (M9)
		"Don't worry about that thing [the scorecard]. Looking daily at the data alone can't provide an accurate view of your work. Good supervisors know that it's more than just numbers and they will tell you when you are doing a good job or not." (M11)
		"If some work a certain way—like, for safety or for the customer—that's how they should work, regardless of some graph. My supervisor gets that, even if some IT system does not." (M14)
		"I've heard mechanics who have been here longer complain about the new scorecards. But then tell me—why are younger, less-experienced workers showing better performance metrics? [Participant-observer asks why.] Because the supervisors have been too nice to the others." (M12)
		I guess sometimes I didn't know when people had a bad day, and now I can. I can go up to [fellow mechanic's name] and be like, 'what happened yesterday?' And sometimes I can even help...." (M1)
		"Now that we know that [mechanic names] are hitting these numbers, [mechanics] who aren't doing as well can go and ask them how. It's good that low-performers can approach high-performers, especially for new mechanics. Some mechanics struggled to improve because they didn't have a good supervisor who could help them do so—either supervisors don't know or they don't care. Now they can get help from others directly." (M10)

Table B.1. (Continued)

Category	Subcategory	Illustrative quotes
		"Talking to supervisors … [makes a grunting sound] … sometimes it's like you are stuck in their check-the-box exercises. It's a total waste of time. Much better to go straight to the source [points at the high-performers on the chart]." (M6)
		"Some supervisors just don't get it. Either they have forgotten what it's like to be out here or never knew in the first place. It's like talking to a brick wall—a total waste of time. This is better because mechanics can support each other directly." (M8)
		"I like having my team directly get the results when I do. They can respond more immediately to problems and it gives me more time to do other things." (S4)
Moderators	Social comparison orientation/avoidance vs. approach	"No one hears anything about being middle of the good [productive] time … but if your bad [nonproductive] time is high, I think people notice … Hell, I notice when others are high … ." (M3)
		"With these new scorecards, it's definitely best to hide in the middle of the pack. No one notices you if you are and that's the best you can hope for." (M4)
		"When I see all those numbers... I check them just to make sure I'm ok, I'm not falling behind … . We all want to do well, of course, but... well, you know, you just don't want to look like you're slacking off or having troubles or stuff like that." (M4)
		"Some [mechanics] tease low-performers and high-performers. If you're low, it's like you're not pulling your weight. If you're high, then you had an easy job or could've helped someone else. It's better off to conform, not excel." (M16)
		[Participant Observer asks how to use the dashboards] "Those... I just glance at them to make sure I'm ok and not screwing up... they're nice to have, but we talk about how things are going later with [supervisor] and the others, so the reports are just a f-up check." (M7)
		"As a former mechanic, I know there are good days and bad days—a single chart can't show that. I've always told my team that if someone wanted to see raw numbers, they could come see me. Giving them the daily data just distracts them from doing their work." (S2)
		"A couple of the mechanics seem to talk about nonproductive time now that we have those new reports [dashboards]." (S6)
Moderators	Intrinsic vs. Extrinsic Motivation	"I like it. It helps us all get better and know each other better—we know who we can learn from, who needs training, what ways there are to improve this job." (M12)
		"Yeah, we're on the road a lot, but we all—you know—work together to help the customers. Gas is important. So we work hard." (M9)
		"I guess I want to avoid being noticed for being too far below others." (M4)
		"When it was just [supervisor name], it was different—he got me, he knows what it's like, that the good is more important than the bad—but the numbers don't, so just be sure the bad doesn't stand out." (M3)
		"I like not being in the middle anymore. Because my team sees what I see, they feel more immediate ownership over the results and believe me when I say I'm not hiding anything." (S3)
		"Even though I would talk to them about it when there was something concerning in my daily reports, now that they can access it directly, [those mechanics] seem to take it more seriously." (S6)
		"Certain mechanics on my team have been difficult to motivate since I started this position. I don't know why, but now they finally seem to be paying attention to how they manage their time—particularly their nonproductive time." (S8)

Appendix C. Pre-experimental Survey of Mechanics

On a scale of 1–4 (1 = never or almost never true of me, 4 = always or almost always true of me), please indicate the extent to which each item describes you.

	Never			Always
Q1. I am strongly motivated by the recognition I can earn from other people	1	2	3	4
Q2. I want other people to find out how good I really can be at my work	1	2	3	4
Q3. To me, success means doing better than other people	1	2	3	4
Q4. I am keenly aware of the promotion goals I have for myself	1	2	3	4
Q5. I am keenly aware of the income goals I have for myself	1	2	3	4
Q6. I enjoy tackling problems that are completely new to me	1	2	3	4
Q7. I enjoy trying to solve complex problems	1	2	3	4
Q8. The more difficult the problem, the more I enjoy trying to solve it	1	2	3	4
Q9. What matters most to me is enjoying what I do	1	2	3	4
Q10. It is important for me to be able to do what I most enjoy	1	2	3	4

Most people compare themselves from time to time with others. For example, they may compare the way they feel, their opinions, their abilities, and/or their situation with those of other people. There is nothing particularly "good" or "bad" about this type of comparison, and some people do it more than others. We would like to find out how often you compare yourself with other people. On a scale of 1–5 (1 = strongly disagree, 5 = strongly agree), please state how much you agree or disagree with the following statements about you.

	Strongly Disagree				Strongly Agree
Q11. I often compare how my loved ones are doing with how others are doing	1	2	3	4	5
Q12. I always pay a lot of attention to how I do things compared with how others do things	1	2	3	4	5
Q13. If I want to find out how well I have done something, I compare what I have done with how others have done	1	2	3	4	5
Q14. I often compare how I am doing socially (e.g., social skills, popularity) with other people	1	2	3	4	5
Q15. I am not the type of person who compares myself often with others	1	2	3	4	5
Q16. I often compare myself with others with respect to what I have accomplished in life	1	2	3	4	5
Q17. I often like to talk with others about mutual opinions and experiences	1	2	3	4	5
Q18. I often try to find out what others think who face problems similar to those I face	1	2	3	4	5
Q19. I always like to know what others in a similar situation would do	1	2	3	4	5
Q20. If I want to learn more about something, I try to find out what others think about it	1	2	3	4	5
Q21. I never consider my situation in life relative to that of other people	1	2	3	4	5

On a scale of 1–5 (1 = almost never to 5 = almost always), please indicate the extent to which each item describes the support provided to you by your immediate supervisor.

	Never				Always
Q22. My work supervisor values my contribution to the company's well-being.	1	2	3	4	5
Q23. My work supervisor strongly considers my goals and values.	1	2	3	4	5
Q24. My work supervisor really cares about my well-being.	1	2	3	4	5
Q25. My work supervisor is willing to help me when I need a special favor.	1	2	3	4	5
Q26. My work supervisor shows very little concern for me.	1	2	3	4	5
Q27. My work supervisor takes pride in my accomplishments at work.	1	2	3	4	5

We would like to ask you about your own assessment of your work.

Please answer the following question (by sliding the button):

	Highest Performer	Lowest Performer
Q28. On "% of my day spent performing productive activities," compared to all 123 mechanics at GasCo, I think my performance would rank me ___ out of 123		

Table D.1. Variations in Treatment Effects Based on Access Frequency

	More views			Fewer views		
	Nonproductive time	Productive time	Support time	Nonproductive time	Productive time	Support time
Treat × *Post*	−3.853**	7.224**	−3.371	−2.743*	−3.802	6.544
Post	1.349	−3.589***	2.240	1.3486	−3.589***	2.240
Individual fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,855	9,855	9,855	9,366	9,366	9,366
Adj. R^2	0.303	0.276	0.156	0.330	0.298	0.163

Notes. This table reports OLS regression results of % *Nonproductive Time* (Columns 1 and 4), % *Productive Time* (Columns 2 and 5), and % *Support Time* (Columns 3 and 6) on a postintervention indicator (*Post*) and an interaction of the two variables (*Treat* × *Post*). *Treat* is absorbed by individual fixed effects and not reported. All variables are defined in Table 1. Standard errors are clustered at the work-center level.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Appendix D. Variations in Treatment Effects Based on Access Frequency to Scorecards

At some treatment centers, more employees paid attention to the scorecards; at others, fewer did so. To estimate treatment effects for these two subsamples, in the table below, we divide the total number of treatment work centers into those with more views (>175 over 65 days) and those with fewer views. We find greater effects on reducing nonproductive time *and* increasing productive time, both in magnitude and statistical significance, in the subsample with more views. The fact that employees who were more likely to view the scorecards ended up experiencing stronger treatment effects is encouraging, as it suggests that the observed changes in behavior were driven by direct data access rather than other concurrent/confounding factors. Furthermore, these results suggest that it is appropriate to focus on our main sample of employees (those who responded to the pre-intervention survey) as they had a stronger intent to view the performance data. The survey and the report cards both required employees to pay attention to their emails and log into their corporate account. Those who responded to the survey were hence much more likely to view their report cards.

Although this set of results yields interesting insights into the average treatment effects, we only include it here as an additional analysis, because the measure for access frequency is at the work center level, not the individual level (our unit of analysis in the study), and viewing the scorecard was a self-selected behavior. Caution therefore needs to be applied to the interpretation of these results.

Endnotes

[1] In our setting, like many professional settings, an employee's time during a workday is automatically measured and tracked in three categories: productive time (spent on revenue-generating activities), support time (spent on work activities that do not directly generate revenue), and nonproductive time (spent on non-work activities). An employee has considerable control over how to allocate time across the three categories (without targets), as more revenue-generating tasks can always be allocated from the queue of customer requests to whomever is ready to take on more work. See Dependent Variables in the Methods section for more detail.

[2] The company, in close partnership with the unions, had designed all the metrics to be automatically tracked both to reduce the burden of collecting the data (i.e., no self-reporting) and to minimize the possibility of gaming.

[3] We borrow the notion that traditional managers are intermediaries from the seminal work of Chandler (1977) on the emergence of managerialism in the late 19th and early 20th centuries, in which he argues that technology in the form of "increased speed and regularity of transportation and communication brought to an end [the] long and expensive chain of [external] middlemen" (p. 214), replacing them with internal middlemen who served the same "administrator" functions (Simon 1947, pp. 39, 326). Across the widely divergent academic disciplines that have found value in studying intermediation, the three common functions through which intermediaries generate and derive value for their clients are strikingly consistent (Spulber 1999): (1) economies of scale and scope—the intermediary, by *aggregating* supply and demand, reduces transaction costs and pools/diversifies risk over time; (2) coordination economies—the intermediary, by *filtering* for fit ("matching") between supply and demand, is recognized as a central, expert place of exchange, thus reducing costs of search and supporting commitment through reliable monitoring and delegation (Chan 1983); and (3) longevity and incentives to build reputation—the intermediary, by *framing* each interaction as long-term and repeated, is recognized as credible by both sides (Diamond 1984), thus enabling it to certify quality (alleviating adverse selection) and make commitments worthy of reliance (mitigating moral hazard). In this paper, we adapt these three functions to the context of supervisor-employee relationships.

[4] The imbalance between the number of employees allocated to the treatment and control groups was due to our agreement with the company and the unions that our intervention would only directly affect a certain number of employees. Union leaders and company managers did not inform employees that some work centers implemented this "pilot program" (the intervention). We compared key employee characteristics between the treatment and control groups and found no differences of statistical and economic significance in age, gender, or race. We did find differences in tenure and preintervention performance metrics (those in the treatment group were about 1.5 years less experienced on average, had about 7.5% less nonproductive time, and 7% more productive time). As a result, we explicitly control for these individual differences in our empirical models.

[5] For qualitative evidence of these conversations in the treatment group, see Table B.1; for the control group, with whom we had less interaction, one mechanic told our participant-observer in a passing conversation: "We get feedback once or twice a day from our bosses … when we are all in the [work center] … at lunch, at the end of the day, sometimes before we start our shift the next day. They know how we did and share that with us. [Participant-Observer: What else do you talk about?] Oh, lots of things … the game last night, the weird food that [mechanic name] brought for lunch, you know … (laughs)."

[6] If the trip took less than 20 minutes, then the actual number of minutes spent on the trip would be counted as productive time.

[7] Performance misjudgment—overconfidence or underconfidence in one's self-assessment of relative performance—is known as the Dunning-Kruger effect (Kruger and Dunning 1999, Dunning 2011). Research has found high-performers underestimating their performance and low-performers overestimating theirs (Burson et al. 2006, Ehrlinger et al. 2008, Schlösser et al. 2013).

[8] Given a relatively small number of clusters (11) in our setting, the cluster-robust standard errors could be downward biased (although "small" is subjective; 10–20 clusters are typically considered "moderate"). Following Cameron et al. (2008), we started with cluster-robust standard errors and then found that our results were robust to their bootstrap-t procedure (the wild cluster bootstrap that corrects for potential downward biases).

[9] Mechanic quotations are demarcated with an M; supervisor quotations are demarcated with an S; and quotations that the participant-observer overheard from a group of mechanics or supervisors are demarcated with a PO. The numbers that follow M or S anonymously but consistently identify a particular individual throughout the manuscript.

[10] The effect of the intervention on % *Productive Time* in column 5 was positive (a 4.5% increase compared with the preintervention mean value) but statistically insignificant ($\beta = 2.697$, $p = 0.438$). Similarly, the effect on % *Support Time* (column 6) was positive but statistically insignificant ($\beta = 0.700$, $p = 0.792$).

References

Aarons-Mele M (2020) Leading through anxiety. *Harvard Bus. Rev.* (May 11), https://hbr.org/2020/05/leading-through-anxiety.

Abuhamdeh S, Csikszentmihalyi M (2009) Intrinsic and extrinsic motivational orientations in the competitive context: An examination of person–situation interactions. *J. Personality* 77(5):1615–1635.

Adams SM (2005) Positive affect and feedback-giving behavior. *J. Management Psych.* 20(1):24–42.

Ahn J, Khandelwal AK, Wei SJ (2011) The role of intermediaries in facilitating trade. *J. Internat. Econom.* 84(1):73–85.

Alder GS (2007) Examining the relationship between feedback and performance in a monitored environment: A clarification and extension of feedback intervention theory. *J. High Tech. Management Res.* 17(2):157–174.

Alder GS, Ambrose ML (2005) An examination of the effect of computerized performance monitoring feedback on monitoring fairness, performance, and satisfaction. *Organ. Behav. Human Decision Processes* 97(2):161–177.

Alvero AM, Bucklin BR, Austin J (2001) An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). *J. Organ. Behav. Management* 21(1):3–29.

Amabile TM, Hill KG, Hennessey BA, Tighe EM (1994) The work preference inventory: Assessing intrinsic and extrinsic motivational orientations. *J. Personality Soc. Psych.* 66(5):950–967.

Anderson S, Rodin J (1989) Is bad news always bad?: Cue and feedback effects on intrinsic motivation. *J. Appl. Soc. Pysch.* 19(6):449–467.

Anon (2015) *Oxford English Dictionary* (Oxford University Press, Oxford, UK).

Ashford SJ, Cummings LL (1983) Feedback as an individual resource: Personal strategies of creating information. *Organ. Behav. Human Performance* 32(3):370–398.

Ashkanani AM, Dunford BB, Mumford KJ (2022) Impact of motivation and workload on service time components: An empirical analysis of call center operations. *Management Sci.* 68(9):6697–6715.

Azmat G, Iriberri N (2016) The provision of relative performance feedback: An analysis of performance and satisfaction. *J. Econom. Management Strategy* 25(1):77–110.

Balcazar F, Hopkins BL, Suarez Y (1985) A critical, objective review of performance feedback. *J. Organ. Behav. Management* 7(3–4):65–89.

Barends E, Rousseau D, Wietrak E, Cioca I (2022) *Performance Feedback: An Evidence Review, Scientific Summary* (Chartered Institute of Personnel and Development, London).

Barnett NP, Murphy JG, Colby SM, Monti PM (2007) Efficacy of counselor vs. computer-delivered intervention with mandated college students. *Addictive Behaviors* 32(11):2529–2548.

Baron RM, Kenny DA (1986) The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J. Personality Soc. Psych.* 51(6):1173–1182.

Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD (2001) Bad is stronger than good. *Rev. General Psych.* 5(4):323–370.

Berman O, Larson RC, Pinker E (1997) Scheduling workforce and workflow in a high volume factory. *Management Sci.* 43(2):158–172.

Bernstein ES (2017) Making transparency transparent: The evolution of observation in management theory. *Acad. Management Ann.* 11(1):1–50.

Bernstein ES (2022) Leveling the "Flatter" playing field. *J. Organ. Des.* 11(1):23–25.

Bernstein ES, Blunden H (2015) The sales director who turned work into a fantasy sports competition. *Harvard Bus. Rev.* (March 27), https://hbr.org/2015/03/the-sales-director-who-turned-work-into-a-fantasy-sports-competition.

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? *Quart. J. Econom.* 119(1):249–275.

Beshears J, Choi JJ, Laibson D, Madrian BC, Milkman KL (2015) The effect of providing peer information on retirement savings decisions. *J. Finance* 70(3):1161–1201.

Blanes I Vidal J, Nossol M (2011) Tournaments without prizes: Evidence from personnel records. *Management Sci.* 57(10):1721–1736.

Bol JC (2008) Subjectivity in compensation contracting. *J. Accounting Literature* 27:1–32.

Bordoloi S, Fitzsimmons J, Fitzsimmons MJ (2019) *Service Management Operations, Strategy, Information Technology* (McGraw-Hill Education, New York).

Botelho TL, Gertsberg M (2022) The disciplining effect of status: Evaluator status awards and observed gender bias in evaluations. *Management Sci.* 68(7):5311–5329.

Brewer N (1995) The effects of monitoring individual and group performance on the distribution of effort across tasks. *J. Appl. Soc. Pysch.* 25(9):760–777.

Burson KA, Larrick RP, Klayman J (2006) Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *J. Personality Soc. Psych.* 90(1): 60–77.

Buunk AP, Gibbons FX (2007) Social comparison: The end of a theory and the emergence of a field. *Organ. Behav. Human Decision Processing* 102(1):3–21.

Cameron LD (2022) "Making Out" while driving: Relational and efficiency games in the gig economy. *Organ. Sci.* 33(1):231–252.

Cameron J, Pierce WD (1994) Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Rev. Ed. Res.* 64(3):363–423.

Cameron J, Pierce WD (1996) The debate about rewards and intrinsic motivation: Protests and accusations do not alter the results. *Rev. Ed. Res.* 66(1):39–51.

Cameron AC, Gelbach JB, Miller DL (2008) Bootstrap-based improvements for inference with clustered errors. *Rev. Econom. Statist.* 90(3):414–427.

Campbell I, Charlesworth S (2012) Salaried lawyers and billable hours: A new perspective from the sociology of work. *Internat. J. Legal Profession* 19(1):89–122.

Cappelli P, Tavis A (2016) The performance management revolution. *Harvard Bus. Rev.* 94(10):58–67.

Carey KB, Scott-Sheldon LAJ, Elliott JC, Garey L, Carey MP (2012) Face-to-face versus computer-delivered alcohol interventions for college drinkers: A meta-analytic review, 1998 to 2010. *Clinical Psych. Rev.* 32(8):690–703.

Carpenter DL, Gregg SR, Owens DS, Buchman TG, Coopersmith CM (2012) Patient-care time allocation by nurse practitioners and physician assistants in the intensive care unit. *Critical Care* 16(1):R27.

Cases-Arce P, Lourenço SM, Martinez-Jerez FA (2017) The performance effect of feedback frequency and detail: Evidence from a field experiment in customer satisfaction. *J. Accounting Res.* 55(5):1051–1088.

Chan YS (1983) On the positive role of financial intermediation in allocation of venture capital in a market with imperfect information. *J. Finance* 38(5):1543–1568.

Chandler AD (1977) *The Visible Hand: The Managerial Revolution in American Business* (Harvard University Press, Cambridge, MA).

Chanen J (2005) The amazing race. *Aba J.* 91(8):46–51.

Chiappa GD, Lorenzo-Romero C, Constantinides E (2014) Disintermediation and user-generated content: A latent segmentation analysis. *Proc. Soc. Behav. Sci.* 148:524–532.

Chircu AM, Kauffman RJ (1999) Strategies for Internet middlemen in the intermediation/disintermediation/reintermediation cycle. *Electronic Marketing* 9(1–2):109–117.

Choi JN (2004) Individual and contextual predictors of creative performance: The mediating role of psychological processes. *Creative Res. J.* 16(2 and 3):187–199.

Christin A (2018) Counting clicks: Quantification and variation in web journalism in the United States and France. *Amer. J. Sociol.* 123(5):1382–1415.

Cialdini RB, Trost MR (1998) Social influence: Social norms, conformity and compliance. Gilbert DT, Fiske ST, Lindzey G, eds. *Handbook of Social Psychology* (McGraw-Hill, New York), 151–192.

Cole HA, Prassel HB, Carlson CR (2018) A meta-analysis of computer-delivered drinking interventions for college students: A comprehensive review of studies from 2010 to 2016. *J. Stud. Alcohol Drugs* 79(5):686–696.

Corbalan G, Paas F, Cuypers H (2010) Computer-based feedback in linear algebra: Effects on transfer performance and motivation. *Comput. Ed.* 55(2):692–703.

Cunningham L, McGregor J (2015) Why big business is falling out of love with the annual performance review. *Washington Post* (August 17), https://www.washingtonpost.com/news/on-leadership/wp/2015/08/17/why-big-business-is-falling-out-of-love-with-annual-performance-reviews/.

DePasque S, Tricomi E (2015) Effects of intrinsic motivation on feedback processing during learning. *Neuroimage* 119:175–186.

Diamond DW (1984) Financial intermediation and delegated monitoring. *Rev. Econom. Stud.* 51(3):393–414.

Donohue B, Allen DN, Maurer A, Ozols J, DeStefano G (2004) Alcohol prevention: A controlled evaluation of two prevention programs in reducing alcohol use among college students at low and high risk for alcohol-related problems. *J. Alcohol Drug Ed.* 48(1):13–33.

Dunn J, Ruedy NE, Schweitzer ME (2012) It hurts both ways: How social comparisons harm affective and cognitive trust. *Organ. Behav. Human Decision Processing* 117(1):2–14.

Dunning D (2011) The Dunning–Kruger effect: On being ignorant of one's own ignorance. *Advances in Experimental Social Psychology* (Elsevier, Amsterdam), 247–296.

Earley PC (1988) Computer-generated performance feedback in the magazine-subscription industry. *Organ. Behav. Human Decision Processing* 41(1):50–64.

Eberly MB, Holley EC, Johnson MD, Mitchell TR (2011) Beyond internal and external: A dyadic theory of relational attributions. *Acad. Management Rev.* 36(4):731–753.

Eberly MB, Holley EC, Johnson MD, Mitchell TR (2017) It's not me, it's not you, it's us! An empirical examination of relational attributions. *J. Appl. Psych.* 102(5):711–731.

Ehrlinger J, Johnson K, Banner M, Dunning D, Kruger J (2008) Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organ. Behav. Human Decision Processing* 105(1):98–121.

Eisenberger R, Stinglhamber F, Vandenberghe C, Sucharski IL, Rhoades L (2002) Perceived supervisor support: Contributions to perceived organizational support and employee retention. *J. Appl. Psych.* 87(3):565–573.

Elliot AJ, Church MA (1997) A hierarchical model of approach and avoidance achievement motivation. *J. Personality Soc. Psych.* 72(1):218–232.

Elliot AJ, Thrash TM (2002) Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *J. Personality Soc. Psych.* 82(5):804–818.

Elliott JC, Carey KB, Bolles JR (2008) Computer-based interventions for college drinking: A qualitative review. *Addiction Behav.* 33(8):994–1005.

Emerson RM, Fretz RI, Shaw LL (1995) *Writing Ethnographic Fieldnotes* (University of Chicago Press, Chicago).

Eren T (2018) Drilling time follow-up with non-productive time monitoring. *Internat. J. Oil Gas Coal Tech.* 19(2):197–216.

Eriksson T, Poulsen A, Villeval MC (2009) Feedback and incentives: Experimental evidence. *Labour Econom.* 16(6):679–688.

Evans JA, Kunda G, Barley SR (2004) Beach time, bridge time, and billable hours: The temporal structure of technical contracting. *Admin. Sci. Quart.* 49:1–38.

Ewenstein B, Hancock B, Komm A (2016) Ahead of the curve: The future of performance management. *The McKinsey Quarterly* (McKinsey & Company, Inc.), 1–10.

Eyring H, Narayanan V (2018) Performance effects of setting a high reference point for peer-performance comparison. *J. Accounting Res.* 56(2):581–615.

Eysenbach G (2008) Medicine 2.0: Social networking, collaboration, participation, apomediation, and openness. *J. Medical Internet Res.* 10(3):1–13.

Fang L, Ivashina V, Lerner J (2015) The disintermediation of financial markets: Direct investing in private equity. *J. Financial Econom.* 116(1):160–178.

Festinger L (1954) A theory of social comparison processes. *Human Relations* 7(2):117–140.

Feys M, Libbrecht N, Anseel F, Lievens F (2008) A closer look at the relationship between justice perceptions and feedback reactions: The role of the quality of the relationship with the supervisor. *Psych. Belgium* 48(2–3):127–156.

Fischer C, Malycha CP, Schafmann E (2019) The influence of intrinsic motivation and synergistic extrinsic motivators on creativity and innovation. *Frontiers Psych.* 10(137):1–10.

Frick B, Simmons R (2008) The impact of managerial quality on organizational performance: Evidence from German soccer. *Management Decision Econom.* 29(7):593–600.

Gagné M, Forest J, Gilbert MH, Aubé C, Morin E, Malorni A (2010) The motivation at work scale: Validation evidence in two languages. *Ed. Psych. Measures* 70(4):628–646.

Gibbons FX, Buunk BP (1999) Individual differences in social comparison: Development of a scale of social comparison orientation. *J. Personality Soc. Psych.* 76(1):129–142.

Gibbs M, Merchant KA, Stede WAV, der, Vargus ME (2004) Determinants and effects of subjectivity in incentives. *Accounting Rev.* 79(2):409–436.

Gilbreth FB, Gilbreth LM (1919) *Applied Motion Study: A Collection of Papers on the Efficient Method to Industrial Preparedness* (The Macmillan Company, New York).

Gilley JW, Gilley AM (2007) *The Manager as Coach* (Greenwood Publishing Group, Westport, CT).

Gittleman M, Pierce B (2013) An improved measure of interindustry pay differentials. *J. Econom. Social Measurement* 38: 229–242.

Goler L, Gale J, Grant A (2016) Let's not kill performance evaluations yet. *Harvard Bus. Rev.* 94(11):90–94.

Guay F, Vallerand RJ, Blanchard C (2000) On the assessment of situational intrinsic and extrinsic motivation: The situational motivation scale (SIMS). *Motivation Emotions* 24(3):175–213.

Guidry G, Spezia K, Salmon G (2018) Collaborative real-time analysis to reduce non-productive time. *SPE Annual Tech. Conf. Exhibition (Dallas, Texas)*.

Haas JR, Hayes SC (2006) When knowing you are doing well hinders performance: Exploring the interaction between rules and feedback. *J. Organ. Behav. Management* 26(1–2):91–111.

Hamel G (2011) First, let's fire all the managers. *Harvard Bus. Rev.* 89(12):48–59.

Hannan RL, Krishnan R, Newman AH (2008) The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *Accounting Rev.* 83(4):893–913.

Hannan RL, McPhee GP, Newman AH, Tafkov ID (2013) The effect of relative performance information on performance and effort allocation in a multi-task environment. *Accounting Rev.* 88(2):553–575.

Harvey JF, Green P (2022) Constructive feedback: When leader agreeableness stifles team reflexivity. *Personality Individual Differences* 194:1–10.

Hebert BG, Vorauer JD (2003) Seeing through the screen: Is evaluative feedback communicated more effectively in face-to-face or computer-mediated exchanges? *Comput. Human Behav.* 19(1):25–38.

Hellervik LW, Hazucha JF, Schneider RJ (1992) Behavior change: Models, methods, and a review of evidence. Dunnette MD, Hough LM, eds. *Handbook of Industrial Organanization Psychology*, 2nd ed. (Consulting Psychologists Press, Palo Alto, CA), 821–895.

Hennessey BA, Zbikowski SM (1993) Immunizing children against the negative effects of reward: A further examination of intrinsic motivation training techniques. *Creative Res. J.* 6(3):297–307.

Hennessey BA, Amabile TM, Martinage M (1989) Immunizing children against the negative effects of reward. *Contemporary Ed. Psych.* 14(3):212–227.

Hersel MC (2020) Fighting with the fourth estate: A theoretical framework of organization-media rivalry for narrative control following a transgression. *Acad. Management Rev.* 47(3):425–447.

Holmstrom B, Milgrom P (1991) Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *J. Law Econom. Organ.* 7:24–52.

Howells J (2006) Intermediation and the role of intermediaries in innovation. *Res. Policy* 35(5):715–728.

Hurt C (2007) Initial public offerings and the failed promise of disintermediation. *Entrepreneurial Bus J.* 2:703–742.

Ilgen DR, Fisher CD, Taylor MS (1979) Consequences of individual feedback on behavior in organizations. *J. Appl. Psych.* 64(4):349–371.

Ilgen DR, Peterson RB, Martin BA, Boeschen DA (1981) Supervisor and subordinate reactions to performance appraisal sessions. *Organ. Behav. Human Performance* 28(3):311–330.

Ivancevich JM, McMahon JT (1982) The effects of goal setting, external feedback, and self-generated feedback on outcome variables: A field experiment. *Acad. Management J.* 25(2):359–372.

Jaccard J, Turrisi R (2003) *Interaction Effects in Multiple Regression* (SAGE, Thousand Oaks, CA).

Jackson E (2012) Ten biggest mistakes bosses make in performance reviews. *Forbes* (January 9), https://www.forbes.com/sites/ericjackson/2012/01/09/ten-reasons-performance-reviews-are-done-terribly/.

Jallat F, Capek MJ (2001) Disintermediation in question: New economy, new networks, new middlemen. *Bus. Horizons* 44(2):55–60.

Johnson DA, Rocheleau JM, Tilka RE (2015) Considerations in feedback delivery: The role of accuracy and type of evaluation. *J. Organ. Behav. Management* 35(3–4):240–258.

Jokisaari M, Nurmi JE (2009) Change in newcomers' supervisor support and socialization outcomes after organizational entry. *Acad. Management J.* 52(3):527–544.

Karabenick SA, Knapp JR (1988) Effects of computer privacy on help-seeking. *J. Appl. Soc. Psych.* 18(6):461–472.

Kc DS, Terwiesch C (2009) Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Management Sci.* 55(9):1486–1498.

Kellogg KC, Valentine MA, Christin A (2020) Algorithms at work: The new contested terrain of control. *Acad. Management Ann.* 14(1):366–410.

Kennedy P (2003) *A Guide to Econometrics* (MIT Press, Cambridge, MA).

Kippenbrock T, Lo WJ, Emory J, Buron B, Odell E, Reimers J (2018) Nurse practitioners' time on nonbillable activities. *J. Amer. Assoc. Nurse Practitioners* 30(9):480–490.

Kluger AN, Adler S (1993) Person-versus computer-mediated feedback. *Comput. Human. Behav.* 9(1):1–16.

Kluger AN, DeNisi A (1996) The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psych. Bull.* 119(2): 254–284.

Kluger AN, Lewinsohn S, Aiello JR (1994) The influence of feedback on mood: Linear effects on pleasantness and curvilinear effects on arousal. *Organ. Behav. Human Decision Processing* 60(2):276–299.

Kraut AI, Pedigo PR, McKenna DD, Dunnette MD (1989) The role of the manager: What's really important in different management jobs. *Acad. Management Executive* 3(4):286–293.

Kruger J, Dunning D (1999) Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J. Personality Soc. Psych.* 77(6):1121–1134.

Kuhnen CM, Tymula A (2012) Feedback, self-esteem, and performance in organizations. *Management Sci.* 58(1):94–113.

Kurtzberg TR, Belkin LY, Naquin CE (2006) The effect of e-mail on attitudes towards performance feedback. *Internat. J. Organ. Anal.* 14(1):4–21.

Kuvaas B (2006) Performance appraisal satisfaction and employee outcomes: Mediating and moderating roles of work motivation. *Internat. J. Human Resources Management* 17(3):504–522.

Larkin I, Pierce L, Gino F (2012) The psychological costs of pay-for-performance: Implications for the strategic compensation of employees. *Strategic Management J.* 33(10):1194–1214.

Larson JR (1986) Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organ. Behav. Human Decision Processing* 37(3):391–408.

Larson JR (1989) The dynamic interplay between employees' feedback-seeking strategies and supervisors' delivery of performance feedback. *Acad. Management Rev.* 14(3):408–422.

Lee MK, Kusbit D, Metsky E, Dabbish L (2015) Working with machines: The impact of algorithmic and data-driven management on human workers. *Proc. 33rd Annual ACM Conf. Human Factors Comput. Systems* (ACM, New York), 1603–1612.

Leicht K, Fennell ML (2001) *Professional Work: A Sociological Approach* (Blackwell Publishers, Oxford, UK).

Leonardi PM, Neeley TB (2022) *The Digital Mindset: What It Really Takes to Thrive in the Age of Data, Algorithms, and AI* (Harvard Business Review Press, Boston, MA).

Lepper MR, Henderlong J (2000) Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. *Intrinsic Extrinsic Motivation and Search Optimization Motivated Performance* (Academic Press, San Diego), 257–307.

London M (2003) *Job Feedback: Giving, Seeking, and Using Feedback for Performance Improvement* (Psychology Press, London).

London M, Smither JW (1995) Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personality Psych.* 48(4):803–839.

Lourenço SM (2016) Monetary incentives, feedback, and recognition—Complements or substitutes? Evidence from a field experiment in a retail services company. *Accounting Rev.* 91(1):279–297.

Lourenço SM, Greenberg JO, Littlefield M, Bates DW, Narayanan V (2018) The performance effect of feedback in a context of negative incentives: Evidence from a field experiment. *Management Accounting Res.* 40:1–14.

Madhavan P, Wiegmann DA (2007) Effects of information source, pedigree, and reliability on operator interaction with decision support systems. *Human Factors* 49(5):773–785.

Maister DH (1993) *Managing the Professional Service Firm* (Simon and Schuster, New York).

Mayer JD, Faber MA, Xu X (2007) Seventy-five years of motivation measures (1930–2005): A descriptive analysis. *Motivation Emotions* 31(2):83–103.

Mayo E (1933) *The Human Problems of an Industrial Civilization* (Macmillan Company, New York).

Mazmanian M, Beckman CM (2018) Making" your numbers: Engendering organizational control through a ritual of quantification. *Organ. Sci.* 29(3):357–379.

McCubbrey DJ, Taylor RG (2005) Disintermediation and reintermediation in the US air travel distribution industry: A Delphi reprise. *Comm. Assoc. Inform. Systems* 15(26):464–477.

McFarland C, Miller DT (1994) The framing of relative performance feedback: Seeing the glass as half empty or half full. *J. Personality Soc. Psych.* 66(6):1061–1073.

Meyer BD (1995) Natural and quasi-experiments in economics. *J. Bus. Econom. Statist.* 13(2):151–161.

Mintzberg H (1975) The manager's job: Folklore and fact. *Harvard Bus. Rev.* 53(4):49–61.

Mintzberg H (2013) *Simply Managing: What Managers Do—And Can Do Better* (Berrett-Koehler Publishers, Oakland, CA).

Mollick ER, Rothbard N (2013) Mandatory fun: Consent, gamification and the impact of games at work. Preprint, submitted June 10, http://dx.doi.org/10.2139/ssrn.2277103.

Murphy P (2010) Web-based collaborative reading exercises for learners in remote locations: The effects of computer-mediated feedback and interaction via computer-mediated communication. *ReCALL* 22(2):112–134.

Nargesi MN (2021) Comparing AI and human feedback: The role of source in feedback perceptions. San Francisco State University. https://doi.org/10.46569/20.500.12680/hd76s531w.

Neighbors C, Larimer ME, Lewis MA (2004) Targeting misperceptions of descriptive drinking norms: Efficacy of a computer-delivered personalized normative feedback intervention. *J. Consulting Clinical Psych.* 72(3):434–447.

Neri A, Cucchiarini C, Strik H (2008) The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2 Dutch. *ReCALL* 20(2):225–243.

Northcraft GB, Ashford SJ (1990) The preservation of self in everyday life: The effects of performance expectations and feedback context on feedback inquiry. *Organ. Behav. Human Decision Processing* 47(1):42–64.

Northcraft GB, Earley PC (1989) Technology, credibility, and feedback use. *Organ. Behav. Human Decision Processes* 44(1):83–96.

Patil SV, Bernstein ES (2022) Uncovering the mitigating psychological response to monitoring technologies: Police body cameras not only constrain but also depolarize. *Organ Sci.* 33(2):541–570.

Peifer C, Schönfeld P, Wolters G, Aust F, Margraf J (2020) Well done! Effects of positive feedback on perceived self-efficacy, flow and performance in a mental arithmetic task. *Frontiers Psych.* 11(1008):1–11.

Pierce L, Snow D, McAfee A (2015) Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Sci.* 61(10):2299–2319.

Pinedo ML (2012) *Scheduling: Theory, Algorithms, and Systems* (Springer, New York).

Pitts FH (2022) Measuring and managing creative labour: Value struggles and billable hours in the creative industries. *Organization* 29(6):1081–1098.

Podsakoff PM, Farh JL (1989) Effects of feedback sign and credibility on goal setting and task performance. *Organ. Behav. Human Decision Processing* 44(1):45–67.

Prabhu V, Sutton C, Sauser W (2008) Creativity and certain personality traits: Understanding the mediating effect of intrinsic motivation. *Creative Res. J.* 20(1):53–66.

Prendergast C (1999) The provision of incentives in firms. *J. Econom. Literature* 37(1):7–63.

Pritchard RD, Bigby DG, Beiting M, Coverdale S, Morgan C (1981) Enhancing productivity through feedback and goal setting. (AFHRL-TR-81-7, AD-A102 032). Manpower and Personnel Division, Air Force Human Resources Laboratory, Brooks AFB, TX.

Puranam P, Alexy O, Reitzig M (2014) What's "new" about new forms of organizing? *Acad. Management Rev.* 39(2):162–180.

Ranganathan A, Benson A (2020) A numbers game: Quantification of work, auto-gamification, and worker productivity. *Amer. Sociol. Rev.* 85(4):573–609.

Reitzig M (2022) *Get Better at Flatter* (Palgrave Macmillan, Cham, Switzerland).

Roberti R, Bartolini E, Mingozzi A (2014) The fixed charge transportation problem: An exact algorithm based on a new integer programming formulation. *Management Sci.* 61(6): 1275–1291.

Roberts JA, Hann IH, Slaughter SA (2006) Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects. *Management Sci.* 52(7):984–999.

Robinson GF, Switzer GE, Cohen ED, Primack BA, Kapoor WN, Seltzer DL, Rubio DM (2014) Shortening the work preference inventory for use with physician scientists: WPI-10. *Clinical Translation Sci.* 7(4):324–328.

Roethlisberger FJ, Dickson WJ (1939) *Management and the Worker: An Account of a Research Program Conducted by the Western Electric Company* (Harvard University Press, Cambridge, MA).

Roth S, Cohen LJ (1986) Approach, avoidance, and coping with stress. *Amer. Psych.* 41(7):813–819.

Schlösser T, Dunning D, Johnson KL, Kruger J (2013) How unaware are the unskilled? Empirical tests of the "signal extraction" counterexplanation for the Dunning–Kruger effect in self-evaluation of performance. *J. Econom. Psych.* 39:85–100.

Schyns B, Schilling J (2013) How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *Leadership Quart.* 24(1):138–158.

Scott K (2017) *Radical Candor: How to Get What You Want by Saying What You Mean* (St. Martin's Press, New York).

Seifert CF, Yukl G, McDonald RA (2003) Effects of multisource feedback and a feedback facilitator on the influence behavior of managers toward subordinates. *J. Appl. Psych.* 88(3): 561–569.

Shanock LR, Eisenberger R (2006) When supervisors feel supported: Relationships with subordinates' perceived supervisor support, perceived organizational support, and performance. *J. Appl. Psych.* 91(3):689–695.

Simon HA (1947) *Administrative Behavior: A Study of Decision Making Processes in Administrative Organization* (Macmillan, New York).

Smith C (2002) The wholesale and retail markets of London, 1660–1840. *Econom. History Rev.* 55(1):31–50.

Snyder RA, Williams RR, Cashman JF (1984) Age, tenure, and work perceptions as predictors of reactions to performance feedback. *J. Psych.* 116:11–21.

Song H, Tucker AL, Murrell KL, Vinson DR (2018) Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. *Management Sci.* 64(6):2628–2649.

Spulber DF (1999) *Market Microstructure: Intermediaries and the Theory of the Firm* (Cambridge University Press, Cambridge, UK).

Staats BR (2018) *Never Stop Learning: Stay Relevant, Reinvent Yourself, and Thrive* (Harvard Business Press, Boston).

Staats BR, Gino F (2012) Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Sci.* 58(6): 1141–1159.

Staats BR, Dai H, Hofmann D, Milkman KL (2017) Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Sci.* 63(5):1563–1585.

Stanton CT, Thomas C (2015) Landing the first job: The value of intermediaries in online hiring. *Rev. Econom. Stud.* 83(2):810–854.

Steelman LA, Wolfeld L (2018) The manager as coach: The role of feedback orientation. *J. Bus. Psych.* 33(1):41–53.

Stock JH, Watson MW (2003) *Introduction to Econometrics* (Pearson Education, New York).

Stone DL, Deadrick DL, Lukaszewski KM, Johnson R (2015) The influence of technology on the future of human resource management. *Human Resources Management Rev.* 25(2):216–231.

Suls J, Martin R, Wheeler L (2002) Social comparison: Why, with whom, and with what effect? *Current Directive Psych. Sci.* 11(5):159–163.

Swift V, Peterson JB (2018) Improving the effectiveness of performance feedback by considering personality traits and task demands. *PLOS One* 13(5):1–18.

Tafkov ID (2013) Private and public relative performance information under different compensation contracts. *Accounting Rev.* 88(1):327–350.

Tapscott D, Tapscott A (2017) How blockchain will change organizations. *MIT Sloan Management Rev.* 58(2):10–13.

Tapscott D, Ticoll D (2003) *The Naked Corporation: How the Age of Transparency Will Revolutionize Business* (Free Press, New York).

Taylor FW (1914) *The Principles of Scientific Management* (Harper and Brothers, New York).

Taylor MS, Fisher CD, Ilgen DR (1984) Individual's reactions to performance feedback in organizations: A control theory perspective. Rowland KM, Ferris GR, eds. *Research in Personnel and Human Resources Management,* vol. 2 (JAI Press, Greenwich, CT), 81–124.

Thornock TA (2016) How the timing of performance feedback impacts individual performance. *Accounting Organ. Soc.* 55:1–11.

Thuillard S, Adams M, Jelmini G, Schmutz S, Sonderegger A, Sauer J (2022) When humans and computers induce social stress through negative feedback: Effects on performance and subjective state. *Comput. Human Behav.* 133:107270.

Tremblay MA, Blanchard CM, Taylor S, Pelletier LG, Villeneuve M (2009) Work Extrinsic and Intrinsic Motivation Scale: Its value for organizational psychology research. *Canadian J. Behav. Sci. Rev.* 41(4):213–226.

Van Dijk D, Kluger AN (2011) Task type as a moderator of positive/negative feedback effects on motivation and performance: A regulatory focus perspective. *J. Organ. Behav.* 32(8): 1084–1105.

Van der Kleij FM, Feskens RCW, Eggen TJHM (2015) Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Rev. Educational Res.* 85(4):475–511.

Weitzel JA, Bernhardt JM, Usdan S, Mays D, Glanz K (2007) Using wireless handheld computers and tailored text messaging to reduce negative consequences of drinking alcohol. *J. Stud. Alcohol Drugs* 68(4):534–537.

Wood JV (1989) Theory and research concerning social comparisons of personal attributes. *Psych. Bull.* 106(2):231–248.

Yakura EK (2001) Billables: The valorization of time in consulting. *Amer. Behav. Sci.* 44(7):1076–1095.

York PL, Prichard DM, Dodson JK, Dodson T, Rosenberg SM, Gala D, Utama B (2009) Eliminating non-productive time associated with drilling through trouble zones. Paper presented at the Offshore Technology Conference, Houston.

Zhang Y, Li H (2010) Innovation search of new ventures in a technology cluster: The role of ties with service intermediaries. *Strategic Management J.* 31(1):88–109.