Effects of Transparent Performance Data on Employee Performance:

Evidence from a Field Experiment[1]

Abstract

There is a growing trend of continuously tracking performance metrics and providing them to employees via digital means without supervisor intermediation. Using a field experiment at a service organization, we examine how employees respond to transparent performance data previously available only to supervisors (i.e., daily performance metrics of employees in the same work group). We find that, compared with the pre-intervention mean value, the treatment group experienced an 11-percent decrease in strictly nonproductive time relative to the control group. The effect on reducing strictly nonproductive time seems greater than that on increasing strictly productive time. Performance improvements are greater in certain employee subsamples: those who previously perceived their supervisors as less-supportive, those with low intrinsic motivation, and those with high extrinsic motivation. We find inconclusive evidence on the moderating effects of social comparison orientation, suggesting that the main effect is unlikely to be driven by access to relative performance information.

[1] The researchers obtained human subjects research approval from their university to conduct this study.

I. INTRODUCTION

Performance feedback has long been an important management control tool to improve employee performance (Lourenco 2016, Bol 2008, Gibbs et al. 2003, Prendergast 1999). Traditionally, employees received such feedback from their supervisors, who carefully aggregated, filtered, and framed it to maximize positive effects on performance (Kluger and DeNisi 1996).

However, with advances in technology, such supervisor-led performance feedback is losing ground in many companies (Cappelli and Tavis 2016)—including over 10 percent of the *Fortune* 500 (Cunningham and McGregor 2015, Ewenstein et al. 2016)—to performance data shared directly with employees via dashboards or other digital means (Goler et al. 2016), reducing or eliminating the information transparency gap between supervisors and employees. Part of a larger trend towards deeper quantification of employee behaviors (Mazmanian and Beckman 2018, Pierce et al. 2015, Ranganathan and Benson 2020) and more immediate reporting of them (Mollick and Rothbard 2014, Staats et al. 2017), employees' access to transparent performance data (i.e., more detailed, more frequent performance data accessible to a broader audience) allows both the employee and the supervisor to receive the same performance information simultaneously, disintermediating supervisors in the performance feedback process.

Using a field experiment, we examine how employees respond to receiving access to transparent performance data previously available only to supervisors. On the one hand, Feedback Intervention Theory (hereafter, FIT; Kluger and DeNisi 1996) indicates that performance feedback is likely to improve performance when it is perceived to be "objective" (e.g., delivered by computers), contains specific task-level information about a familiar job, or is delivered at higher frequency, because such feedback focuses employee attention towards the task-motivation and task-learning levels of their cognitive processes. Negative effects from adding subjectivity to performance data have similarly been documented by the performance evaluation literature in accounting and economics (Bol 2008, Gibbs et al. 2003, Prendergast 1999) because it leads to biases (such as leniency bias and centrality bias), increases favoritism, and causes inconsistencies across supervisors and across employees. Access to transparent performance data without supervisor intermediation should, to a large extent, (a) focus employees' efforts on improving task performance, (b) reduce the downside of managerial subjectivity, and (c) improve performance.

On the other hand, when supervisors are intermediaries, they aggregate, filter, and frame the performance information to (a) minimize the information-processing costs for employees, (b) enable employees to interpret the data, (c) focus employees' attention on the most relevant information, and (d) offer actionable guidance. When performance data flows directly to employees without supervisor intermediation, they may experience greater difficulty finding, recognizing, making sense of, and using it. Akin to Casas-Arce et al.'s (2017) finding that more (and more frequent) feedback may not improve employee performance, access to the same performance data as supervisors receive may not help—and could even hurt—performance.

The purpose of this paper is therefore to address the open—and important—empirical question of whether or not access to transparent performance data, not intermediated by supervisors, improves employee performance. To do so, we conducted a field experiment with embedded-participant observation at a large US gas utility. 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 employees: supervisors could see daily performance metrics from all employees under their supervision—automatically tracked by sensors in employees' trucks, computers, and handheld devices—which the employees themselves could not see until supervisors chose to share it. During our study, a randomly selected group of frontline employees ("service mechanics") received access to transparent performance data—that is, direct access to individual-level performance metrics. These were exactly the same data that had previously been available only to their supervisors. Over several months, we compared this treatment group's performance with that of the control group that did not receive direct access to these data. In a difference-in-differences analysis, we find that, compared with the pre-intervention mean value, the treatment group experienced an 11-percent decrease in strictly nonproductive time relative to the control group, suggesting that providing access to transparent performance data improved employee performance.[2] We find a greater treatment effect on reducing strictly nonproductive time than on increasing strictly productive time (in both magnitude and significance), suggesting that employees—left to make sense of performance data without supervisor intermediation—focused more on reducing activities perceived as "bad" or "nonproductive" than on increasing activities perceived as "good" or "productive" (Baumeister et al. 2001, p. 323).

Two moderating effects allow us to unpack that result. Consistent with our view of transparent performance data as potential substitutes for "bad" supervisors, we find that employees who perceived low supervisor support *ex ante* had treatment effects of greater magnitude and statistical significance for both reducing strictly nonproductive time and increasing strictly productive time. Consistent with our view of transparent performance data as presenting greater external performance pressure and information-processing challenges, we find that access to such data impacts employees with different motivation types in predictably different ways: performance improvement effects are greater for those with relatively low intrinsic motivation or relatively high extrinsic motivation.

[2] In our setting, an employee's time during a workday is automatically measured and tracked in three categories: strictly productive time (spent on revenue-generating activities), support time (spent on work activities that do not directly generate revenue), and strictly nonproductive time (spent on non-work activities). An employee has considerable control over how to allocate time across the three categories, as more revenue-generating tasks can always be allocated from the queue of customer requests to whomever is ready to take on more work.

Given prior research on relative performance information, we expected that our treatment—which likely intensified the access to relative performance information already available through supervisor-led performance feedback—might improve performance due to even greater access to relative performance information (Tafkov 2013, Kuhnen and Tymula 2012). We therefore expected greater treatment effects for employees with stronger social comparison orientation. However, we found mixed and inconclusive evidence: those with *weak* social comparison orientation reduced strictly nonproductive time more while those with strong social comparison orientation seemed to have increased strictly productive time more, although the latter effect is statistically insignificant. Our main effect of access to transparent performance data therefore seems unlikely to be primarily driven by greater access to relative performance information.

These results were reinforced by the qualitative evidence collected by our embedded participant observer, which provides additional insights into how employees and supervisors reacted to the intervention. Employees appreciated seeing performance data whenever they wanted. They spoke about not "standing out" for their nonproductive time rather than about improving their productive time. And those previously dissatisfied with their supervisors welcomed access to transparent performance data with more enthusiasm and perceived greater benefit from it.

This paper makes several contributions. First and foremost, we contribute to the literature in accounting and management on performance feedback. While a vast literature on the effects of performance feedback has clarified the relationship between performance improvement and certain attributes of the feedback (such as frequency, specificity, and negative or positive framing), certain attributes of the recipient (such as gender and past performance), or whether and how the feedback is linked to incentives or targets (Eyring and Narayanan 2018, Lourenço et al. 2018, Thornock 2016, Hannan et al. 2008), none directly examines the effect of the current trend towards eliminating the information transparency gap between employees and supervisors. To address that theory-practice gap, in this study, we conceptualize employees' access to transparent performance data as a form of disintermediation of supervisors and find that directly providing the system-generated performance data increased individual performance. By looking at the data sources and the data intermediaries separately, we see that interactions with data sources (rather than data intermediaries) may shift attention toward avoiding behaviors perceived to be "bad" (nonproductive) rather than toward behaviors perceived to be "good" (productive).

Second, we contribute to the literature in management control on nonmonetary motivating mechanisms. While most of that research studies the effects of performance feedback either directly linked to an incentive or in a setting with high-powered incentives, the vast majority of employees do not have high-powered incentives (Gibbs et al. 2003, Holmstrom and Milgrom 1991). By studying a low-powered–incentive setting, we show that information by itself can powerfully affect employee behaviors. While a nascent set of accounting studies has also begun to examine the effects of information disclosure alone on behavioral and performance changes, particularly in the context of relative performance feedback (Eyring and Narayanan 2018), our study differs in that we compare the effect of performance information with and without supervisor intermediation. By focusing on the role of supervisors when feedback is provided in a low-powered–incentive environment, we are able to show that employees' relationships with supervisors play a key moderating role in the effects of direct access to the same performance data that supervisors have—a finding that can considerably advance that literature.

Third, we contribute to the emerging literature on the use of digital information systems to improve performance through better knowledge-sharing on performance-relevant tasks or decisions (Li and Sandino 2018). By showing that more transparent sharing of performance data can have powerful behavioral effects and by showing that those behavioral effects depend on the "quality" of the supervisors, we provide greater insight into the question of whether technology and data act best as a substitute for bad supervisors or as a complement to good ones. By examining individual work motivation type, we also provide insights into the type of employees such digital information-sharing systems are most likely to help.

This study has practical implications for the growing trend of continuously tracking detailed performance metrics and providing them to employees via digital means without supervisor intermediation. Our results suggest that organizations considering investing in information systems to do so should first take into account (a) behavioral tendencies—such as a focus on bad behaviors over good ones, (b) relational considerations—such as the quality of relationships with supervisors, and (c) individual work motivation types—such as intrinsic and extrinsic motivation.

The rest of the paper is organized as follows: Section II presents our hypotheses development. Section III describes the research setting, method, and data. Section IV discusses the analyses and findings. Section V concludes.

II. THEORY AND HYPOTHESES DEVELOPMENT

Manager-led Performance Feedback

Organizations have long used performance feedback—information about the effectiveness of an employee's work behavior (Ashford and Cummings 1983, Taylor et al. 1984)—to provide knowledge of results (Ammons 1956), "cue" efforts to improve those results (Vroom 1964), and improve employee productivity (Alvero et al. 2001, Balcazar et al. 1985, Ilgen et al. 1979, Larson 1989). Studies in psychology, organizational behavior, and operations demonstrate the potentially beneficial effects of manager-intermediated feedback on performance (e.g., Ilgen et al. 1979, Ivancevich and McMahon 1982, Pritchard et al. 1981). However, empirical research also demonstrates that performance feedback does not always improve performance (Lourenço 2016, Eriksson et al. 2009); 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). Such mixed findings led to further studies of how such heterogeneous effects depend on the characteristics of the feedback (Casas-Arce et al. 2017), the incentive scheme (Hannan et al. 2008, Tafkov 2013), and the individual's position in the performance distribution (Eyring and Narayanan 2018; Lourenço et al. 2018).

Most notably, Kluger and DeNisi's (1996) review of empirical studies, accounting for 607 effect sizes and 23,663 observations, found that while feedback interventions did improve performance on average, 38 percent of them decreased it. Based on this meta-analysis of empirical studies, Kluger and DeNisi (1996) developed the Feedback Intervention Theory (FIT). FIT explains that the variance between feedback's "positive" and "negative" performance outcomes is based on how it focuses an employee's locus of 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 to manage evaluations of the self and their implications). Using past empirical studies, Kluger and DeNisi test FIT's propositions against a long list of moderators and find that managerial cues directing attention to meta-task processes make feedback less effective, while those directing attention to task-motivation or task-learning processes make it more effective. The over 4,000 works that have cited FIT over the past 20+ years make a good case that it offers a comprehensive framework with which to evaluate empirical evidence on how performance feedback interventions affect performance in organizations. Specifically, comparing how organizations provide performance feedback to the conditions described in FIT can help to better predict the feedback's likely performance effects.

In practice, supervisors (direct managers) have traditionally played a central, intermediating role in crafting feedback to meet FIT's propositions. Supervisors typically receive more detailed and more timely performance information about all 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 daily performance data to provide 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. In addition to this intermediating role,[3] good supervisors can provide coaching, motivation, or action plans to help employees improve. It would then seem to follow that good supervisors—i.e., those who deliver performance feedback as FIT suggests it is most optimally delivered—would drive the best outcomes from performance feedback interventions. That has been the perspective of many organizations as they have sought better outcomes from performance feedback systems.

Yet as traditional middlemen continue to be disintermediated in many areas by technology,[4] do supervisors now face the same risk in one of their core functions—as intermediaries in performance feedback processes? What would happen if the same performance data, previously available only to the supervisors, became transparent to employees?

Effects of Access to Transparent Performance Data

FIT predicts the conditions under which performance feedback can improve performance:

> Specifically, an FI [feedback intervention] provided for a familiar task, containing cues that support learning, attracting attention to feedback-standard discrepancies at the task level

[3] We borrow the notion that managers are intermediaries from Chandler's (1977) seminal work 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" (Chandler 1977, p. 214), which was replaced by internal middlemen who served the same "administrator" functions (Simon 1947, pp. 39, 326).
[4] Such disintermediation has been observed in financial markets (Fang et al. 2015, Morrison 2005), supply chains (Jallat and Capek 2001, Mills and Camek 2004, Waldfogel and Reimers 2015), innovation markets (De Silva et al. 2018, Howells 2006), governance bodies (Ahn et al. 2011, Gellman 1996, Iansiti and Lakhani 2017), and more (for a review, see Chircu and Kauffman 1999).

> (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 1 [standard deviation]. (Kluger and DeNisi 1996, p. 278)

FIT therefore suggests at least two reasons to expect better performance from employees given access to performance data previously available only to supervisors. First, the disintermediated data will be closer to the task level and will be available at a higher frequency ("velocity FI"), likely directing employees' attention to task learning and task motivation. Second, having the data delivered directly, without supervisor intermediation, is likely to decrease "cues to the meta-task level"; that is, employees are less likely to feel evaluative pressure and to distract themselves by thinking about what the supervisor's view of their performance means for them (rather than about how to improve that performance).

On the other hand, supervisors could add other types of value to the traditional performance feedback process—value which might be lost in bypassing that process. Employees facing much more performance data may struggle to find or recognize the most relevant information or to realize what actions would be most appropriate. Access to information about how one's own performance compares to that of others working for the same supervisor—information to which previously only the supervisor had access—could lead to distraction.

It is thus an empirical question whether access to transparent performance data will affect performance positively or negatively. We therefore state the following two-sided hypothesis:

Hypothesis 1a (H1a): Providing employees with access to transparent performance data previously available only to supervisors will increase employee performance.
Hypothesis 1b (H1b): Providing employees with access to transparent performance data previously available only to supervisors will decrease employee performance.

We also recognize the possibility that a simple dichotomy of performance changes may not capture a sufficiently nuanced view of the effect. In testing Hypotheses 1a and 1b, we will therefore examine the impact of transparent performance data on both reducing strictly nonproductive activities and increasing strictly productive activities. This is because transparent performance data might not only affect average performance, but also reallocate effort amongst different activities (Brewer 1995, Hannan et al. 2013). Some work activities are perceived to be strictly good or productive (such as revenue-generating activities), others are perceived to be strictly nonproductive (such as personal phone calls and coffee breaks), and still others are perceived to be "neutral" (necessary work activities, such as training and maintenance, that do not directly generate revenue but make long-term revenue generation possible). Because research has shown many ways in which "bad is stronger than good" (Baumeister et al. 2001, p. 323), we will specifically explore whether making performance across multiple categories of activity transparent—without supervisor intermediation—might naturally focus employee attention more on avoiding activities that could be considered "bad" (strictly nonproductive) than on increasing activities considered to be "good" (strictly productive).

The Moderating Role of Supervisor Support

How will employee performance shift when supervisors no longer hold information privilege and employees no longer exclusively rely on them for feedback? That would seem to depend on the existing relationship between supervisors and employees. Specifically, the degree to which employees perceive their supervisors as supportive (and hence trust them) would seem to affect how much they might respond to an intervention making them less reliant on their supervisors. Perceived supervisor support has been tied to in-role and extra-role performance (Eisenberger et al. 2002, Jokisaari and Nurmi 2009, Shanock and Eisenberger 2006) but, to our knowledge, the effect of that relationship on the effect of transparent performance data remains unstudied.

On the one hand, access to the same performance data that supervisors receive can only convey *what* one's performance is, not *how* to improve it (Hellervik et al. 1992, London and Smither 1995).

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, while 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 the non-informational dimensions of the role, such as providing guidance and inspiration, that would *complement* transparent performance data. 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 performance data may be a *substitute* for the supervisor’s role in performance feedback (Hamel 2011, Tapscott and Ticoll 2003), especially for “bad” supervisors whom employees don’t trust. A key potential benefit of having supervisors intermediating the “raw” performance data is that they can apply subjective assessments that can correct distortions in the objective measures and take into account uncontrollable factors so as to make fair assessments. However, subjectivity also introduces potential biases and increases influencing activities (Bol 2008). Gibbs et al. (2003) find that subjective assessments work best when employees have strong trust in their supervisors. Without such trust, they will likely find supervisors’ subjectivity in performance feedback to be “harmful” and will benefit more from access to transparent performance data without that intermediation. In addition, finding out where you stand in comparison to coworkers can provide you with other role models—high-performing peers (Song et al. 2017)—who may be better coaches than your own supervisor (e.g., Ilgen et al. 1981, Larson 1986). Thus, access to transparent performance data, offering actionable insights by identifying peer role models and "best practices," may make the supervisor's input less valuable than it had been. This also suggests that, once given direct access, employees with less-supportive supervisors are likely to find all the benefits and none of the "harms" they perceived when they received feedback from those supervisors.

One might therefore expect that, in the case of supervisor-intermediated performance feedback, better supervisors—better at supporting their employees' work—are less likely to be replaced by employees' direct access to performance data. Performance gains from access to transparent performance data would then come primarily from employees with supervisors perceived to be less supportive. We therefore hypothesize:

Hypothesis 2 (H2): When employees are provided with transparent performance data previously available only to supervisors, those with previously less-supportive supervisors will improve their performance more than those with previously more-supportive supervisors.

The Moderating Role of Work Motivation Type

Facing transparent performance data (e.g., a digital dashboard showing daily performance metrics for all employees in the same work unit) could generate higher external performance pressure than the less-transparent feedback process. Whether that has positive or negative performance effects would depend on how it changes motivation. Research has shown that 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), others by external inducements such as pay or social recognition (i.e., high extrinsic motivation). Intrinsic and extrinsic motivation orientations are likely to have different moderating effects on performance when employees have access to transparent performance data. Employees who are mainly intrinsically motivated are less likely to be affected by external performance pressures such as transparent performance data, while those who are mainly extrinsically motivated may be more likely to be affected. In other words, transparent performance data could act as a substitute for intrinsic motivation but as a complement to extrinsic motivation. We therefore hypothesize:

Hypothesis 3 (H3): When employees are provided with transparent performance data, those with weaker intrinsic motivation about the work will improve their performance more than those with stronger intrinsic motivation.
Hypothesis 4 (H4): When employees are provided with transparent performance data, those with stronger extrinsic motivation about the work will improve their performance more than those with weaker extrinsic motivation.

The Moderating Role of Social Comparison Orientation

Although employees may already have some access to relative performance information, transparent performance data is likely to make it a more prominent aspect of performance feedback. Festinger's (1954) seminal paper proposed that, in the absence of clear standards of correctness, people evaluate themselves, their opinions, and their capabilities in comparison with those of others (for detailed reviews of social comparison orientation theory, see Gibbons and Buunk (1999), Suls, Martin, and Wheeler (2002), and Wood (1989)). Knowing how one performs relative to others can improve performance (Tafkov 2013, Kuhnen and Tymula 2012, Azmat and Iriberri 2010), in part due to peer pressure mechanisms.

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, using the established Iowa-Netherlands Comparison Orientation Measure (Gibbons and Buunk 1999, p. 199). If the main effect of our study's intervention is driven by greater access to relative performance information and a desire to outperform peers, we expect that those more prone to social comparison will be more likely to improve after gaining such access. We therefore hypothesize:

Hypothesis 5 (H5): When employees are provided with transparent performance data, those with stronger social comparison orientation will improve their performance more than those with weaker social comparison orientation.

III. METHODS AND DATA

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 off gas, repair gas appliances, repair leaks, and respond to emergencies. They typically start their day by logging into the system on their trucks, then reviewing and accepting orders made available by dispatch. A mechanic maps a path to an order, arrives on site, completes the order, then drives to the next order or takes a break. Although the activities may appear routine, mechanics self-identify as and are considered to be 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. This identity as "highly trained professionals" makes intrinsic motivation and social reputation relevant concerns in the presence of transparent performance feedback. 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 criteria 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, 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., if the mechanics in the control group were to learn about the difference in treatment status through interactions with those in the treatment group). 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 the real workplace. In fact, many, if not most, frontline jobs lack strong incentives (see US 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 percent of total compensation. Nor were there strong career-related incentives: because the workforce was unionized, there was little concern about job loss based on individual performance and it was widely 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, ranging 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 by creating a consistent set of metrics. Through a bottom-up effort (including the mechanics and their unions), GasCo had generated a single scorecard of metrics—collected automatically by technology in the mechanics' trucks and computers—to which the mechanics had collectively agreed. There were three umbrella categories of mechanic-specific metrics for performance: percentages of productive time, support time, and nonproductive time. (These are defined in the "Dependent variables" subsection below.)

Data Collection and Measures

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 are 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.[5] 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 only available to supervisors (in the control group where mechanics only had access to performance information via supervisor intermediation).

This intervention was set up and 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 (Bandiera et al. 2011, Floyd and List 2016). Because our study was a natural field experiment, 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."

-------------------- Insert Figure 1 About Here --------------------

-------------------- Insert Figure 2 About Here --------------------

[5] 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).

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 to a mechanic in the treatment group (with names disguised)—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 pre-intervention period to take into account the full cycle of seasonal effects.

To gather moderating and control variables, we sent a survey 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 the researchers, not GasCo, and that no responses would be seen by anyone in the company. The email contained a link to an external Qualtrics survey website. We sent the survey to all mechanics and had a 51-percent response rate, which management reported was typical for this population. Comparing human-resources data on our control variables (tenure, age, and race) for responding and nonresponding mechanics revealed no bias in the type of individual who responded. Table 1 shows the definitions and descriptive statistics for our key variables and Table 2 tabulates the unconditional correlations between them.

-------------------- Insert Table 1 and Table 2 About Here --------------------

Dependent variables. To measure performance, GasCo used three standardized metrics for efficient allocation of time, for which a mechanic had sole accountability: *% Productive Time*, *% Support Time*, and *% Nonproductive Time*, which collectively represent the entire workday. Each category was clearly and stringently defined and tracked (see Figure 2).

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, en route time was constrained based on real-time traffic data, while onsite time was constrained based on calculated times for the work. That is, a trip from here to there counted as 20 minutes of productive time if that's 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 maintenance), 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. While 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 directly tag their time to one category or another in the system; their daily work time was automatically tracked and coded for one of those three categories. To further prevent abuse or gaming, time allocation and activity records were audited. Indeed, one reason GasCo thought greater transparency in disclosing performance data might be beneficial for the company was that it might reduce the possibility for abuse by increasing the number of eyeballs on the data.

GasCo designed these metrics to help mechanics allocate more time to productive activities. Our decision to use these measures of time allocation as measures of performance was therefore

[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.

primarily driven by GasCo's own focus, which therefore deserves elaboration here. GasCo's focus on time allocation reflects the increasing importance for distributed workforces (such as GasCo's mechanics) of making wise use of their time. But we also found support for time allocation as a performance measure in several influential literatures, including operations management literatures on scheduling (Pinedo 2012) in factories (Berman et al. 1997), healthcare (Kc and Terwiesch 2009), trucking (Roberti et al. 2014), and financial services (Staats and Gino 2012). Although terminology varies across industries, past research has—just as GasCo did—grouped employee activities into three categories (e.g., Malos and Campion 2000) generically termed billable, support, and nonbillable (at GasCo: productive, support, and nonproductive, respectively).[7] That is, some activities directly generate revenue, some are work-related but do not directly generate revenue, while others are not work-related at all.

All three GasCo performance metrics—*% Productive Time*, *% Support Time*, and *% Nonproductive Time*—were in use long before our study, as was the automated system which tracked them; the study was solely focused on the effect of making the resulting performance data transparent to mechanics. In our empirical analysis, *% Nonproductive Time* and *% Productive Time* are the dependent variables. Because a workday is allocated entirely between the three categories, showing changes in two is sufficient to fully reflect any change in how mechanics allocate time. Throughout the period of our study, there was enough demand for both productive and support activities to keep all mechanics busy 100 percent 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 percent

[7] While GasCo management viewed "productive" and "nonproductive" time as equivalent to "billable" and "nonbillable" time, prior discussions with the union had resulted in the terms "productive" and "nonproductive" because the end-customer was not directly billed for time.

while the average *% Productive Time* is 59.1 percent. Standard deviation values suggest considerable variation in these dependent variables.

Moderating and control variables. In order to measure perceived supervisor support and other moderating variables, we administered a pre-experimental survey (Appendix 1) to all the mechanics in our sample. Consistent with Eisenberger et al. (2002) and Shanock and Eisenberger (2006), we used the six-item instrument to measure perceived supervisor support (Q22–Q27 in Appendix 1). Our measure for perceived supervisor support, consistent with its design and previous use, 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.

The Work Preference Inventory (WPI), a well-established instrument for measuring motivation at work, has frequently been used to measure employees' intrinsic and extrinsic motivation orientation (Amabile, Hill, Hennessey, and Tighe 1994, Robinson et al. 2014). We used it in our pre-intervention survey. Our measure for intrinsic motivation is the sum of the scores from the five survey questions related to it (Q6–Q10 in Appendix 1), while the measure for extrinsic motivation is the sum of the scores from the five questions related to it (Q1–Q5 in Appendix 1), 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 also asked the mechanics to complete the established scale of social comparison orientation, the Iowa-Netherlands Comparison Orientation Measure (Buunk and Gibbons 2007, Gibbons and Buunk 1999). Our measure for *Social Comparison Orientation*, consistent with its design and previous use, is the sum of the scores from each relevant survey question (Q11–Q21 in Appendix 1), adjusted for reverse-coding. Cronbach's alpha for these questions is 0.8824.

We control for past performance in all regressions, since 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.[8] Our measure of a mechanic's actual past performance comes from archival data (averaged over the pre-intervention period) on *% Nonproductive Time* and *% Productive Time*. Our measure of self-assessed past performance comes from a survey question (Q28 in Appendix 1). We standardized the responses on a scale of 0 to 100. By pairing a mechanic's response with actual pre-intervention performance, we could control for past performance, both actual and perceived (self-assessed) (Meyer 1995).

We also incorporate demographic control variables—including tenure, age, and race—based on GasCo's internal records. On average, the mechanics had worked at the company for over 18.5 years and were 47.2 years old. About 55 percent were white.

From Table 2, we see that performance is negatively correlated with tenure and age (suggesting a decrease in productive efforts with the increase of seniority) and positively correlated with perceived supervisor support and self-evaluation. The moderating variables are moderately correlated. We control for all these variables or individual fixed effects in our regressions. While we do not explicitly control for work quality, GasCo did constantly 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.

Participant observation. To gather qualitative evidence and understand how mechanics and

[8] Performance misjudgment—overconfidence or underconfidence in one's self-assessment of relative performance—is known as the Dunning-Kruger effect (Dunning 2011, Kruger and Dunning 1999). 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).

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 research. For a week, he rode, worked, ate, and hung out with other mechanics as a typical apprentice, rotating amongst work centers. His note-taking 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 was a natural and frequent topic. This qualitative evidence 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.

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:

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

Or

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

where Y_{it} is the performance metric at the employee(i)-workday(t) level, $Treatment_{it}$ is an indicator variable that equals 1 if the employee worked in one of the work centers in the treatment group and $Post_{it}$ is an indicator variable that equals 1 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. 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 our Hypothesis 1 (the average treatment effect), we ran two specifications: one including as *Controls* all the available employee-level control measures (*Tenure*, *Age*, *Supervisor Support*, etc.), 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).

IV. RESULTS

H1a and H1b: Does Access to Transparent Performance Data Improve Performance?

-------------------- Insert Table 3 and Figure 3 About Here --------------------

Table 3 shows the main regression results with and without individual fixed effects. Columns 1 and 3 use *% Nonproductive Time* as the dependent variable, while Columns 2 and 4 use *% Productive Time*. The coefficient on *Treat x Post* is the estimated treatment effect of the intervention on the performance metric. The negative and statistically significant coefficients on this interaction term in Columns 1 and 3 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 pre-intervention mean value, the results in Column 3 indicate an 11-percent decrease in *% Nonproductive Time*, suggesting that access to transparent performance data previously available only to supervisors improved performance (H1a). As one employee told the participant-observer, "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." Appendix 2 provides additional qualitative evidence about how mechanics and supervisors reacted to the intervention.

Columns 2 and 4 of Table 3, however, present an important caveat to the results for H1a. The effect of the intervention on *% Productive Time* was positive (a 4.5-percent increase in Column 4 based on the pre-intervention mean value), but not statistically significant after we clustered the standard errors by work group. Comparing Columns 1 and 3 with 2 and 4, 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, 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. This suggests that access to transparent performance data triggers a greater behavioral shift towards reducing strictly nonproductive behaviors than toward increasing strictly productive behaviors. As perceived by the mechanics and captured by the participant-observer, it made 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" and to "conform, not excel." 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." Appendix 2 provides additional qualitative evidence behind this result. 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.

To ensure the quality of our intervention—that is, to be sure that access to transparent performance data and not something unrelated is driving the results—we obtained the number of views of the scorecards for each treatment site. Across the four sites, the average employee accessed the daily report at least once every three days; some far more frequently. Sites whose mechanics accessed the report most often had the greatest productivity boost, indicating that it was, indeed, transparent performance data that drove the effects.

H2: The Moderating Role of Supervisor Support

To test H2, we ran the baseline regressions, 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 or (b) below the median, following the precedent set by prior studies of perceived supervisor support; these groups are labelled "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 once given 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, consistent with H2 and suggesting that access to transparent performance data (previously available only to supervisors) could serve as a substitute for "bad" supervisors. Compared with the pre-intervention mean values, we see a 16.27-percent decrease in *% Nonproductive Time* and a 16.07-percent increase in *% Productive Time* for the low-supervisor-support subsample. Calculating the z-statistics that compare the coefficients on *Treat x Post* between Columns 3 and

4 (Stock and Watson 2003), the difference in treatment effects for *% Productive Time* between the high- and low-supervisor-support subsamples is statistically significant at the five-percent level. Our participant-observer heard one mechanic say that he liked the fact that “low-performers can approach high-performers”—knowing now who they were—to learn how to improve if they “didn’t have a good supervisor.” Employees with less supervisor support improved more “because mechanics can support each other directly,” whereas, for those employees, discussions with their supervisors had been just “check-the-box exercises” and “a total waste of time.” Appendix 2 provides additional qualitative evidence about 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.” In contrast, transparent performance data were described as “just about the work” and a “reality check that I can trust.” While supervisors overall received strong ratings, mechanics who worked for the less-supportive ones benefited more when those supervisors were disintermediated with transparent performance data, which acted as a surrogate for the supervisors.

-------------------- Insert Table 4 About Here --------------------

H3 and H4: The Moderating Role of Work Motivation Type

To test H3, we ran the baseline regressions, splitting the sample into those who reported intrinsic motivation that was (a) greater than or equal to the sample median or (b) below the median; these groups are labelled “High Intrinsic Motivation” and “Low Intrinsic Motivation” in Table 5. This analysis examines whether mechanics with lower intrinsic motivation for work showed more performance improvement once given 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, consistent with H3 and suggesting that access to transparent performance data (previously available only to supervisors) could serve as a substitute for intrinsic motivation. Compared with the pre-intervention mean values, we see a 21.28-percent decrease in *% Nonproductive Time* and a 13.85-percent increase in *% Productive Time* for the low-intrinsic-motivation subsample. Calculating the z-statistics that compare the coefficients on *Treat x 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 one-percent level. The difference in treatment effects for *% Productive Time* between these two subsamples is statistically significant at the five-percent level.

-------------------- Insert Table 5 About Here --------------------

To test H4, we ran the baseline regressions, splitting the sample into those who reported extrinsic motivation that was (a) greater than or equal to the sample median or (b) below the median; these groups are labelled "High Extrinsic Motivation" and "Low Extrinsic Motivation" in Table 6. This analysis examines whether mechanics with higher extrinsic work motivation showed more performance improvement once given access to transparent performance data.

We find treatment effects of greater magnitude for both *% Nonproductive Time* and *% Productive Time* for those who had high extrinsic motivation, consistent with H4 and suggesting that access to transparent performance data could serve as a complement for extrinsic motivation. Compared with the pre-intervention mean values, we see a 15.20-percent decrease in *% Nonproductive Time* and a 13.32-percent increase in *% Productive Time* for the high-extrinsic-motivation subsample. Calculating the z-statistics that compare the coefficients on *Treat x 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 one-percent level. However, the difference in treatment effects for *% Nonproductive Time* between these two subsamples is statistically insignificant.

-------------------- Insert Table 6 About Here --------------------

H5: The Moderating Role of Social Comparison Orientation

To test H5, we ran the baseline regressions, splitting the sample into those who reported social comparison orientation that was (a) greater than or equal to the sample median or (b) below the median; these groups are labelled "High Social Comparison" and "Low Social Comparison" in Table 7. This analysis examines whether mechanics who have higher social comparison orientation (and hence are more influenced by relative performance information) showed more performance improvement once given access to transparent performance data.

The treatment effect on % *Nonproductive Time* is larger in magnitude for the subsample with *low* social comparison orientation, *inconsistent* with H4. The treatment effect on *% Productive Time* is larger in magnitude for the high-social-comparison-orientation subsample, but the difference in treatment effects is statistically insignificant. We therefore cannot assign social comparison orientation a significant moderating role in the effect of transparent performance data on performance. In fact, Columns 1 and 2 suggest that those with low social comparison orientation benefited more from the intervention, as they reduced more nonproductive time.

There are several possible explanations for these mixed and inconclusive results related to H5. In real work environments (as opposed to a laboratory setting), it's 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 formally could be smaller. Another possibility is that heightened social comparison could lead to negative psychological consequences such as diminished trust in coworkers (Dunn, Ruedy, and Schweitzer 2012), discouragement (Beshears, Choi, Laibson, Madrian, and Milkman 2015), and unproductive behaviors. In any case, the results in Table 7 suggest that the main effect of our intervention is unlikely to be mainly driven by greater access to relative performance information.

-------------------- Insert Table 7 About Here --------------------

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 part, in the significant coefficient on our time variable, *Post*, in some regressions. In part to account for seasonality, we requested a much longer time series of pre-intervention performance data and re-ran the regressions using month fixed effects. The results were similar. Second, because our dependent variables, *% Nonproductive Time* and *% Productive 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. For simpler interpretation, we report OLS with standard errors clustered by work location. Fourth, a mechanic's daily performance data are correlated with past performance. In our reported analysis, we controlled for pre-intervention performance (actual and self-assessed) and clustered standard errors by work location (to account for the correlation between all mechanics' performance data over time 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 also without the self-assessed performance variable. In both cases, results were similar.

V. CONCLUSION

Our goal in this study was to examine how access to transparent performance data that had been available only to supervisors affects employee performance. In a field experiment in a large U.S. service organization, our findings support that, on average, providing access to transparent performance data does increase employee performance. Yet there were two important nuances: First, access to such data did not encourage an increase in the "best" (strictly productive) behaviors as much as it encouraged avoidance of the "worst" (strictly unproductive). Second, the access produced 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.

We identify the causal relationship between employee performance and access to transparent performance data using a field experiment. However, as with any field experiment of this complexity, identifying causal relationships comes at a cost. The single US setting, single-period intervention, single-profession subject pool, and relatively independent work all limit the generalizability of our findings. Future field research can extend our study by exploring how these findings—and particularly the relevant moderators—might change in other cultures, in other professions, and with more interdependent work. In addition, we did not investigate certain other moderators that might make the intervention more successful, such as how supervisors are trained to give feedback in the new system and ways to help the traditional and new systems work well together. Future research can examine these “value-adding” roles supervisors play in the feedback process and how they can continue to provide value even when technology equalizes access to performance data.

Harkening back to the time-use studies of scientific management, our main result prompts us to wonder if this new wrinkle—giving workers themselves access to the data rather than asking supervisors to aggregate, filter, and frame it first—helps as much as the organizations adopting the approach expect. Especially during 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 a transition to transparent performance data might be affecting differently situated employees differently. Our field experiment shows that the performance benefits for an employee depend on how supportive his or her supervisor already is, on his or her work motivation type, and on whether the performance data gives transparency to measures of productive (billable), support or nonproductive (non-billable) activities. 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 be evenly conferred.

REFERENCES

Ahn, J., A. K. Khandelwal, S.-J. Wei. 2011. The role of intermediaries in facilitating trade. *Journal of International Economics* **84**(1) 73–85.

Alvero, A. M., B. R. Bucklin, J. Austin. 2001. An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). *Journal of Organizational Behavior Management* **21**(1) 3–29.

Amabile, T. M., Hill, K. G., Hennessey, B. A., & Tighe, E. M. 1994. The Work Preference Inventory: Assessing intrinsic and extrinsic motivational orientations. *Journal of Personality and Social Psychology* **66**(5), 950–967.

Ammons, R. B. 1956. Effects of knowledge of performance: A survey and tentative theoretical formulation. *Journal of General Psychology* **54**(2) 279–299.

Ashford, S. J., L. L. Cummings. 1983. Feedback as an individual resource: Personal strategies of creating information. *Organizational Behavior and Human Performance* **32**(3) 370–398.

Azmat, G., Iriberri, N., 2010. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics* **94** 7-8.

Balcazar, F., B. L. Hopkins, Y. Suarez. 1985. A critical, objective review of performance feedback. *Journal of Organizational Behavior Management* **7**(3–4) 65–89.

Bandiera, O., Barankay, I., & Rasul, I. 2011. Field experiments with firms. *Journal of Economic Perspectives*, 25(3), 63–82.

Barankay, I., 2012. Rank incentives: Evidence from a randomized workplace experiment. Working paper. The Wharton School of Business, University of Pennsylvania, Philadelphia.

Barankay, I., 2011. Rankings and social tournaments: Evidence from a crowd-sourcing experiment. Working paper, The Wharton School of Business, University of Pennsylvania, Philadelphia.

Baumeister, R. F., E. Bratslavsky, C. Finkenauer, K. D. Vohs. 2001. Bad is stronger than good. *Review of General Psychology* **5**(4) 323–370.

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

Bertrand, M., E. Duflo, S. Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* **119**(1) 249–275.

Bol, J. C. 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature* **27** 1–32.

Brewer, N. 1995. The effects of monitoring individual and group performance on the distribution of effort across tasks. *Journal of Applied Social Psychology* **25**(9) 760–777.

Burson, K. A., R. P. Larrick, J. Klayman. 2006. Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology* **90**(1) 60–77.

Buunk, A. P., F. X. Gibbons. 2007. Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes* **102**(1) 3–21.

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

Casas-Arce, P., S. M. Lourenço, F. A. Martinez-Jerez. 2017. The performance effect of feedback frequency and detail: Evidence from a field experiment in customer satisfaction. *Journal of Accounting Research* **55**(5) 1051–1088.

Casas-Arce, P., Martinez-Jerez, F. A., 2009. Relative performance compensation, contests, and dynamic incentives. *Management Science* **55**(8).

Chandler, A. D. 1977. *The visible hand: The managerial revolution in American business*. Cambridge, MA, Harvard University Press.

Chircu, A. M., R. J. Kauffman. 1999. Strategies for Internet middlemen in the intermediation/disintermediation/reintermediation cycle. *Electronic Markets* **9**(1–2) 109–117.

Christensen, L. T., G. Cheney. 2014. Peering into transparency: Challenging ideals, proxies, and organizational practices. *Communication Theory* **25**(1) 70–90.

Cialdini, R. B., M. R. Trost. 1998. Social influence: Social norms, conformity and compliance. Gilbert, D. T., and Fiske, S. T., and Lindzey, G., eds. *The Handbook of Social Psychology*. New York, McGraw-Hill, 151–192.

Clogg, C. C., E. Petkova, A. Haritou. 1995. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology* **100**(5) 1261–1293.

Cunningham, L., J. McGregor. 2015. Why big business is falling out of love with the annual performance review. *Washington Post* (August 17). Available at: https://www.washingtonpost.com/news/on-leadership/wp/2015/08/17/why-big-business-is-falling-out-of-love-with-annual-performance-reviews/, accessed May 2016.

De Silva, M., J. Howells, M. Meyer. 2018. Innovation intermediaries and collaboration: Knowledge-based practices and internal value creation. *Research Policy* **47**(1) 70–87.

Dunning, D. 2011. The Dunning-Kruger effect: On being ignorant of one's own ignorance. *Advances in Experimental Social Psychology* **44** 247–296.

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

Eisenberger, R., F. Stinglhamber, C. Vandenberghe, I. L. Sucharski, L. Rhoades. 2002. Perceived supervisor support: Contributions to perceived organizational support and employee retention. *Journal of Applied Psychology* **87**(3) 565–573.

Emerson, R. M., R. I. Fretz, L. L. Shaw. 1995. *Writing ethnographic fieldnotes*. Chicago, University of Chicago Press.

Eriksson, T., Poulsen, A., Villeval, M. C., 2009. Feedback and incentives: Experimental evidence. *Labor Economics* **16**(6), 679-688.

Ewenstein, B., B. Hancock, A. Komm. 2016. Ahead of the curve: The future of performance management. *McKinsey Quarterly* (May). Available at: http://www.mckinsey.com/business-functions/organization/our-insights/ahead-of-the-curve-the-future-of-performance-management, accessed May 2016.

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

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

Floyd, E., & List, J. A. 2016. Using Field Experiments in Accounting and Finance. *Journal of Accounting Research*, *54*(2), 437–475.

Gellman, R. 1996. Disintermediation and the Internet. *Government Information Quarterly* **13**(1) 1–8.

Gibbons, F. X., B. P. Buunk. 1999. Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology* **76**(1) 129–142.

Gibbs, M., K. Merchant, W. Van der Stede, M. Vargus. 2004. Determinants and effects of subjectivity in incentives. *The Accounting Review* **79**(2) 409–436.

Gilley, J. W., A. M. Gilley. 2007. *The manager as coach*. Westport, CT, Greenwood.

Gittleman, M., B. Pierce. 2013. How prevalent is performance-related pay in the United States? Current incidence and recent trends. *National Institute Economic Review* **226** R4-R16. doi:10.1177/002795011322600102.

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

Haas, J. R., S. C. Hayes. 2006. When knowing you are doing well hinders performance: Exploring the interaction between rules and feedback. *Journal of Organizational Behavior Management* **26**(1–2) 91–111.

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

Hannan, R. L., G. P. McPhee, A. H. Newman, I. D. Tafkov. 2013. The effect of relative performance information on performance and effort allocation in a multi-task environment. *The Accounting Review* **88**(2) 553–575.

Hannan, R. L., Krishnan, R., Newman, A. H., 2008. The effects of disseminating relative performance feedback in tournament and individual performance compensation plans. *Accounting Review* **83**(4), 893-913.

Hellervik, L. W., J. F. Hazucha, R. J. Schneider. 1992. Behavior change: Models, methods, and a review of evidence. Dunnette, M. D. and Hough, L. M., eds. *Handbook of Industrial and Organizational Psychology*, 2nd ed. Palo Alto, CA, Consulting Psychologists Press, 821–895.

Holmstrom, B., P. Milgrom. 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization* **7** 24–52.

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

Iansiti, M., K. R. Lakhani. 2017. The truth about blockchain. *Harvard Business Review* **95**(1) 118–127.

Ilgen, D. R., C. D. Fisher, M. S. Taylor. 1979. Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology* **64**(4) 349–371.

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

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

Jokisaari, M., J.-E. Nurmi. 2009. Change in newcomers' supervisor support and socialization outcomes after organizational entry. *Academy of Management Journal* **52**(3) 527–544.

Kandel, E., Lazear, E. P., 1992. Peer pressure and partnerships. *Journal of Political Economics* **100** (4), 801-817.

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

Kennedy, P. 2003. *A guide to econometrics*. Cambridge, MA, MIT Press.

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

Kluger, A. N., S. Lewinsohn, J. R. Aiello. 1994. The influence of feedback on mood: Linear effects on pleasantness and curvilinear effects on arousal. *Organizational Behavior and Human Decision Processes* **60**(2) 276–299.

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

Kuhnen, C. M., Tymula, A., 2012. Feedback, self-esteem, and performance in organizations. *Management Science* **58**(1) 94-113.

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

Larson, J. R. 1986. Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organizational Behavior and Human Decision Processes* **37**(3) 391–408.

Larson, J. R. 1989. The dynamic interplay between employees' feedback-seeking strategies and supervisors' delivery of performance feedback. *Academy of Management Review* **14**(3) 408–422.

Li, S. X., T. Sandino. 2018. Effects of an information sharing system on employee creativity, engagement, and performance. *Journal of Accounting Research* **56**(2) 713–747.

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

Lourenço, S. M., Greenberg, J. O., Littlefield, M., Bates, D. W., Narayanan, V. G., 2018. The performance effect of feedback in a context of negative incentives: Evidence from a field experiment. *Management Accounting Research* **40**, 1-14.

Lourenço, S. M., 2016. Monetary Incentives, Feedback, and Recognition—Complements or Substitutes? Evidence from a Field Experiment in a Retail Services Company. *Accounting Review* **91**(1), 279-297.

Malos, S. B., M. A. Campion. 2000. Human resource strategy and career mobility in professional service firms: A test of an options-based model. *Academy of Management Journal* **43**(4) 749–760.

Martinko, M. J., W. L. Gardner. 1982. Learned helplessness: An alternative explanation for performance deficits. *Academy of Management Review* **7**(2) 195–204.

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

Meyer, B. D. 1995. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics* **13**(2) 151–161.

Mills, J. F., V. Camek. 2004. The risks, threats and opportunities of disintermediation: A distributor's view. *International Journal of Physical Distribution & Logistics Management* **34**(9) 714–727.

Mollick, E. R., N. Rothbard. 2014. Mandatory fun: Consent, gamification and the impact of games at work. *The Wharton School Research Paper Series*.

Morrison, A. D. 2005. Credit derivatives, disintermediation, and investment decisions. *The Journal of Business* **78**(2) 621–648.

Northcraft, G. B., S. J. Ashford. 1990. The preservation of self in everyday life: The effects of performance expectations and feedback context on feedback inquiry. *Organizational Behavior and Human Decision Processes* **47**(1) 42–64.

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

Pinedo, M. L. 2012. *Scheduling: Theory, algorithms, and systems*. New York, Springer.

Podsakoff, P. M., J. L. Farh. 1989. Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes* **44**(1) 45–67.

Prendergast, C. 1999. The provision of incentives in firms. *Journal of Economic Literature* **37**(1) 7–63.

Pritchard, R. D., D. G. Bigby, M. Beiting, S. Coverdale, C. Morgan. 1981. Enhancing productivity through feedback and goal setting. *DTIC Document*. Available at: https://apps.dtic.mil/docs/citations/ADA102032, accessed April 2019.

Ranganathan, A., A. Benson. 2020. A numbers game: Quantification of work, auto-gamification, and worker productivity. *American Sociological Review* **85**(4) 573-609.

Riegel, D. 2018. How to solicit negative feedback when your manager doesn't want to give it. *Harvard Business Review*. Available at: https://hbr.org/2018/03/how-to-solicit-negative-feedback-when-your-manager-doesnt-want-to-give-it, accessed January 2021.

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

Rynes, S. L., A. E. Colbert, K. G. Brown. 2002. HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management* **41**(2) 149–174.

Schlösser, T., D. Dunning, K. L. Johnson, J. Kruger. 2013. How unaware are the unskilled? Empirical tests of the "signal extraction" counterexplanation for the Dunning-Kruger effect in self-evaluation of performance. *Journal of Economic Psychology* **39** 85–100.

Seifert, C. F., G. Yukl, R. A. McDonald. 2003. Effects of multisource feedback and a feedback facilitator on the influence behavior of managers toward subordinates. *Journal of Applied Psychology* **88**(3) 561–569.

Shanock, L. R., R. Eisenberger. 2006. When supervisors feel supported: Relationships with subordinates' perceived supervisor support, perceived organizational support, and performance. *Journal of Applied Psychology* **91**(3) 689–695.

Simon, H. A. 1947. *Administrative behavior: A study of decision making processes in administrative organization*. New York, Macmillan.

Song, H., A. L. Tucker, K. L. Murrell, D. R. Vinson. 2017. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices.

Management Science **64**(6) 2628–2649.

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

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

Tafkov, I. D., 2013. Private and public relative performance information under different compensation contracts. *Accounting Review* **88**(1), 327-350.

Tapscott, D., D. Ticoll. 2003. *The naked corporation: How the age of transparency will revolutionize business*. New York, Free Press.

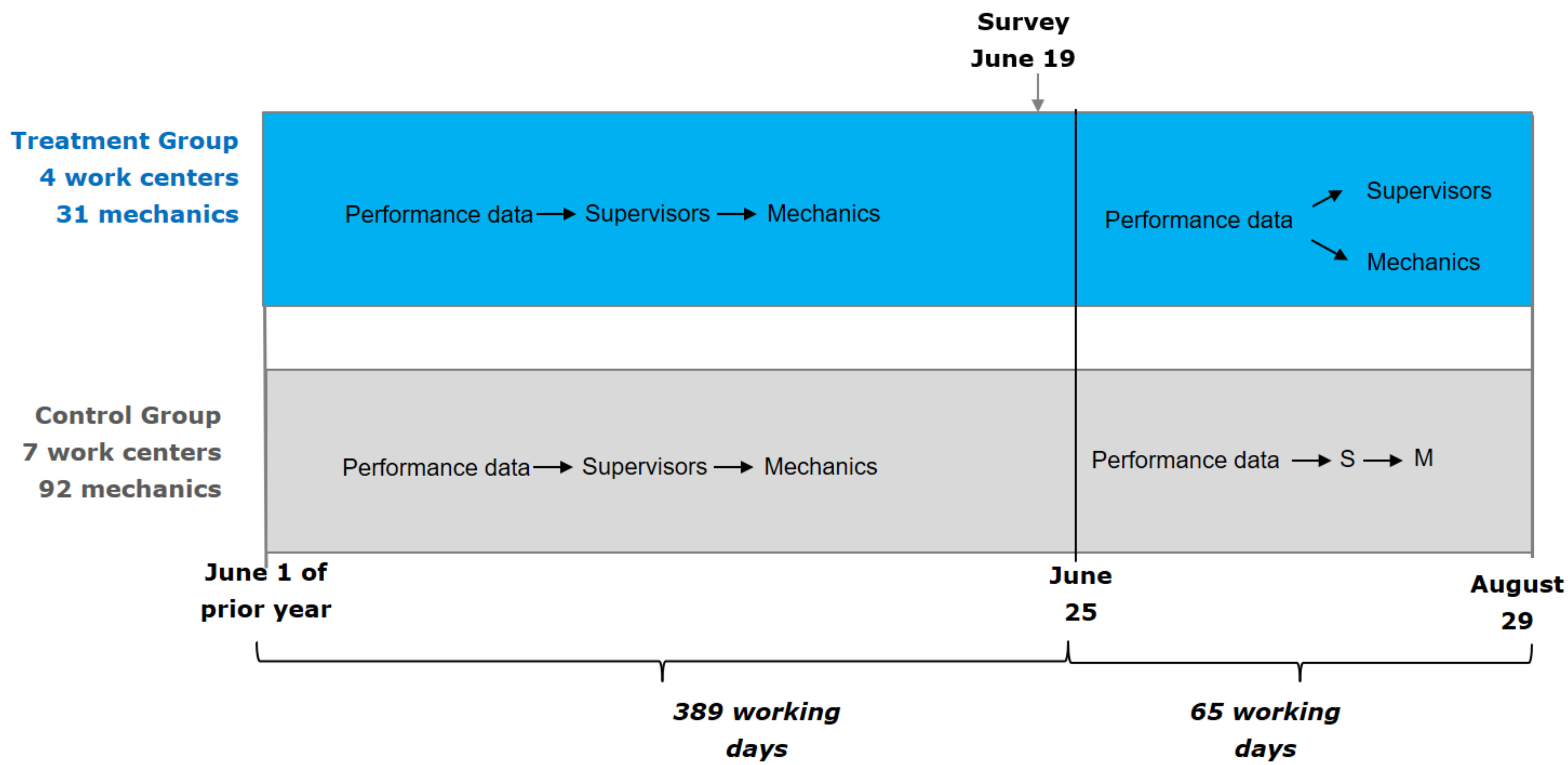
Taylor, M. S., C. D. Fisher, D. R. Ilgen. 1984. Individual's reactions to performance feedback in organizations: A control theory perspective. K. M. Rowland and G. R. Ferris, eds. *Research in Personnel and Human Resources Management*. Greenwich, CT, JAI Press, 81–124.

Thornock, T.A. 2016. How the timing of performance feedback impacts individual performance. *Accounting, Organizations and Society* **55** 1–11.

Vroom, V. H. 1964. *Work and motivation*. New York, Wiley.

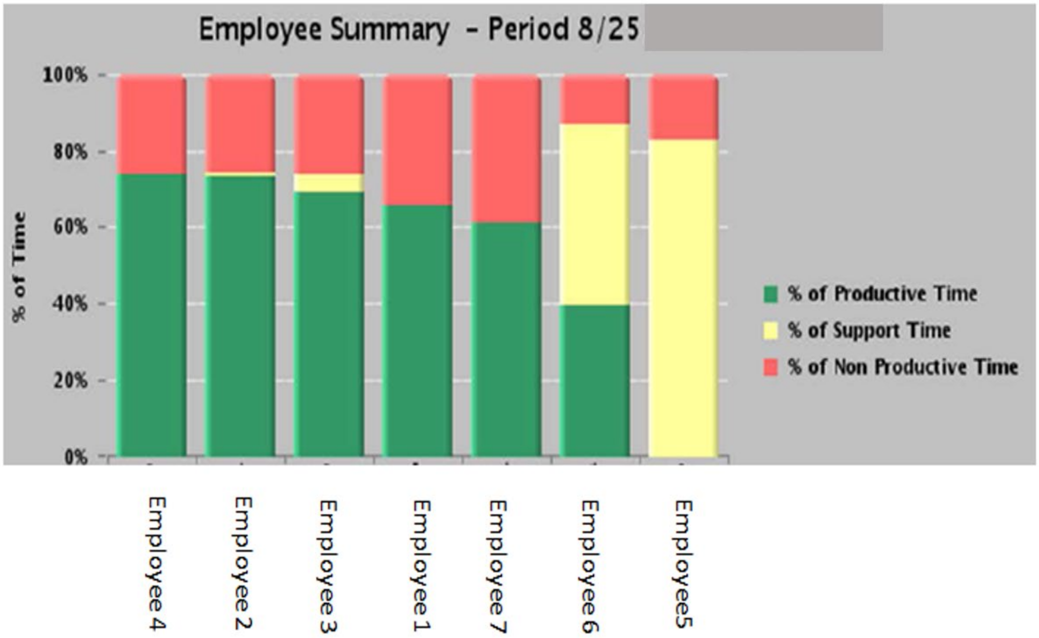
Waldfogel, J., I. Reimers. 2015. Storming the gatekeepers: Digital disintermediation in the market for books. *Information Economics and Policy* **31** 47–58.

Figure 1 Randomization and Timeline



S: Supervisors; M: Mechanics

Figure 2
Sample Screenshot of the Mechanic's View of the Daily Scorecard Information and Time/Activity Categorization



Productive Time

- En route
- Onsite

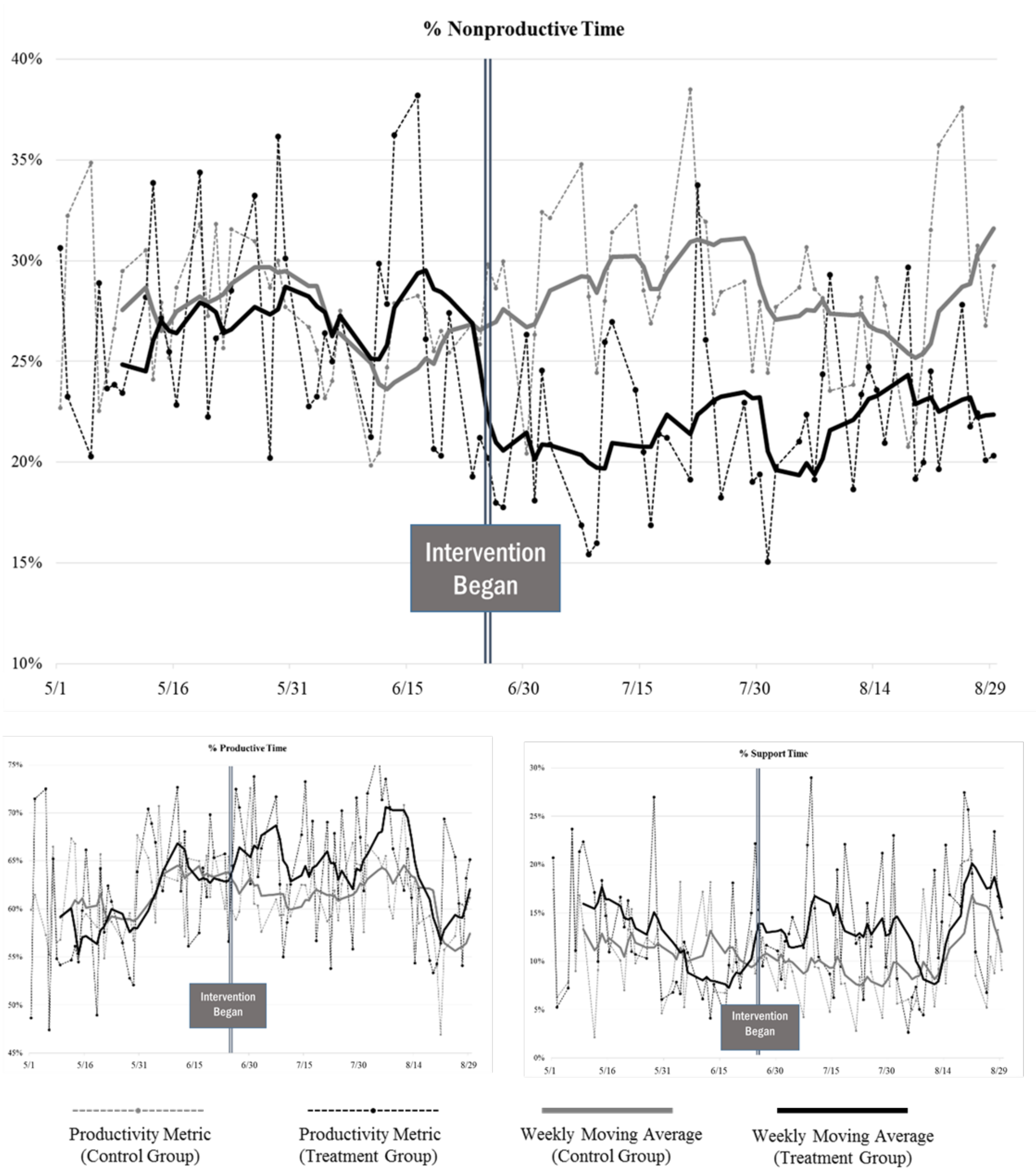
Support Time

- Load/Unload
- Standby
- Building Maintenance
- Vehicle Maintenance
- Meeting
- Training
- Union Business
- Chart Change
- Material Pick up

Non Productive

- Travel (Non Order)
- Lunch
- Break
- Personal
- Ready

Figure 3
Visualization of Treatment Effects



Legend: Individual data points reflect the daily average performance of the treatment or control group on the productivity metric shown in the title of the chart (*% Nonproductive Time*, *% Productive Time*, and *% Support Time*). The dotted lines connect the daily data points to show the variation from one day to the next. The thicker solid lines reflect the weekly moving average of those individual data points.

Table 1
Dependent and Independent Variables

Variable	Definition	Mean	SD	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
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 1).	24.164	4.007	24.000
Social Comparison Orientation	Sum of the scores for questions on social comparison orientation (Q11–Q21 in Appendix 1).	35.193	7.653	37.000
Intrinsic Motivation	Sum of the scores for questions on intrinsic motivation (Q6–Q10 in Appendix 1).	16.527	2.362	16.000
Extrinsic Motivation	Sum of the scores for questions on extrinsic motivation (Q1–Q5 in Appendix 1).	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
Self-evaluation	The self-reported percentile of the pre-intervention performance on *% Productive Time* (Q28 in the Appendix 1).	67.331	31.705	81.890

Notes: n=11,120.

Table 2
Correlations

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

n = 11,120; all correlations in bold are significant at $p < .01$.

Table 3
H1: Does Access to Transparent Performance Data Improve Performance?

	(1) % Non-productive Time	(2) % Productive Time	(3) % Non-productive Time	(4) % Productive Time
Treat x Post	-3.410**	2.846	-3.397**	2.697
	(1.148)	(3.336)	(1.103)	(3.349)
Treat	0.299	-0.081		
	(0.237)	(0.226)		
Post	1.254	-3.528**	1.349	-3.589***
	(1.041)	(0.654)	(1.058)	(0.718)
Tenure	0.023	-0.048**		
	(0.018)	(0.020)		
Age	-0.061**	0.037		
	(0.022)	(0.026)		
White	-0.326	-0.071		
	(0.376)	(0.408)		
Supervisor Support	0.076	-0.148		
	(0.080)	(0.079)		
Social Comparison	-0.072*	0.070**		
	(0.036)	(0.023)		
Intrinsic Motivation	-0.165**	0.152*		
	(0.074)	(0.072)		
Extrinsic Motivation	0.182*	-0.135		
	(0.084)	(0.080)		
Prior Performance (Productive)		0.999***		
		(0.010)		
Prior Performance (Nonproductive)	1.005***			
	(0.006)			
Self-evaluation	-0.007*	-0.004		
	(0.003)	(0.004)		
Individual fixed effects?	No	No	Yes	Yes
Observations	11,120	11,120	11,120	11,120
Adj. rsq.	0.29	0.26	0.29	0.26

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of *% Nonproductive Time* (Columns 1 and 3) and *% Productive Time* (Columns 2 and 4) on a treatment indicator (*Treat*), a post-intervention indicator (*Post*), an interaction of the two variables (*Treat x Post*), and other controls. In Columns 3 and 4, *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 reported in parentheses.

Table 4
H2: Moderating Role of Supervisor Support

	(1) % Non-productive Time High Supervisor Support		(2) % Non-productive Time Low Supervisor Support	(3) % Productive Time High Supervisor Support		(4) % Productive Time Low Supervisor Support
Treat x Post	-2.258	z=-1.34	-5.298**	0.790	z=2.01**	9.377***
	(1.337)		(1.825)	(4.052)		(1.386)
Post	-0.278		3.004**	-3.537**		-3.670**
	(1.278)		(1.177)	(1.453)		(0.892)
Individual fixed effects?	Yes		Yes	Yes		Yes
Observations	7,159		3,961	7,159		3,961
Adj. rsq.	0.24		0.36	0.24		0.31

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of *% Nonproductive Time* (Columns 1–2) and *% Productive Time* (Columns 3–4) on a post-intervention indicator (*Post*), an interaction of the two variables (*Treat x 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 1. Standard errors are clustered at the work-center level and reported in parentheses.

Comparing the coefficients on *Treat x Post* between Columns 1 and 2 (Columns 3 and 4) yields a z-statistic of -1.34 (2.01).

Table 5
H3: Moderating Role of Intrinsic Motivation

	(1) % Non-productive Time High Intrinsic Motivation	**(2) % Non-productive Time Low Intrinsic Motivation**	**(3) % Productive Time High Intrinsic Motivation**	**(4) % Productive Time Low Intrinsic Motivation**
Treat x Post	-0.649 z=-3.13***	-6.711**	-2.003 z=2.14**	8.303***
	(1.694)	(0.939)	(4.318)	(2.116)
Post	0.624	2.258*	-3.287**	-3.968**
	(1.401)	(1.055)	(1.068)	(0.957)
Individual fixed effects?	Yes	Yes	Yes	Yes
Observations	6,050	5,070	6,050	5,070
Adj. rsq.	0.25	0.34	0.29	0.29

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of *% Nonproductive Time* (Columns 1–2) and *% Productive Time* (Columns 3–4) on a post-intervention indicator (*Post*), an interaction of the two variables (*Treat x 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 1. Standard errors are clustered at the work-center level and reported in parentheses.

Comparing the coefficients on *Treat x Post* between Columns 1 and 2 (Columns 3 and 4) yields a z-statistic of -3.13 (2.14).

Table 6
H4 - Moderating Role of Extrinsic Motivation

	(1) % Non-productive Time High Extrinsic Motivation		**(2) % Non-productive Time Low Extrinsic Motivation**	**(3) % Productive Time High Extrinsic Motivation**		**(4) % Productive Time Low Extrinsic Motivation**
Treat x Post	-4.579**	z=-1.01	-2.858**	7.875***	z=1.75*	-0.754
	(1.441)		(0.899)	(1.218)		(4.779)
Post	1.134		1.682*	-3.822**		-3.227
	(1.541)		(0.795)	(0.934)		(0.845)
Individual fixed effects?	Yes		Yes	Yes		Yes
Observations	6,054		5,066	6,054		5,066
Adj. rsq.	0.25		0.34	0.26		0.36

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of *% Nonproductive Time* (Columns 1–2) and *% Productive Time* (Columns 3–4) on a post-intervention indicator (*Post*), an interaction of the two variables (*Treat x 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 1. Standard errors are clustered at the work-center level and reported in parentheses.

Comparing the coefficients on *Treat x Post* between Columns 1 and 2 (Columns 3 and 4) yields a z-statistic of -1.01 (1.75).

Table 7
H5: Moderating Role of Social Comparison Orientation

	(1) % Non-productive Time High Social Comparison		**(2) % Non-productive Time Low Social Comparison**	**(3) % Productive Time High Social Comparison**		**(4) % Productive Time Low Social Comparison**
Treat x Post	-2.549	z=-1.28	-5.526**	4.829	z=0.37	2.506
	(1.875)		(1.369)	(3.371)		(5.203)
Post	-0.302		3.970**	-2.446**		-5.403**
	(1.252)		(1.371)	(1.043)		(1.827)
Individual fixed effects?	Yes		Yes	Yes		Yes
Observations	5,909		5,211	5,909		5,211
Adj. rsq.	0.29		0.29	0.26		0.26

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of *% Nonproductive Time* (Columns 1–2) and *% Productive Time* (Columns 3–4) on a post-intervention indicator (*Post*), an interaction of the two variables (*Treat x 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 1. Standard errors are clustered at the work-center level and reported in parentheses.

Comparing the coefficients on *Treat x Post* between Columns 1 and 2 (Columns 3 and 4) yields a z-statistic of -1.28 (0.37).

Appendix 1: 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		

Appendix 2: Reactions from Mechanics and Supervisors to Providing Employees with Transparent Performance Data (from Participant Observation)

"I finally know what management sees about me! Now I understand some of the feedback I've gotten in the past." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

"No one hears anything about being middle of the good [productive] time… but if your bad [non-productive] 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." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

"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]." [Mechanic]

"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." [Mechanic]

"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." [Mechanic]

“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.” [Supervisor]

“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.” [Supervisor]

“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.” [Supervisor]

“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.” [Supervisor]