Seeing Where You Stand:
From Performance Feedback to Performance Transparency

Ethan S. Bernstein
Harvard Business School
e@hbs.edu

Shelley Xin Li
University of Southern California
xli766@usc.edu

ABSTRACT
Advanced technology is enabling organizations to continuously track employee activities, reporting transparent performance data that increasingly replace traditional manager-led performance feedback reviews. Using a field experiment at a large service organization, we examine the effect of replacing performance feedback reviews with system-generated transparent performance data on employees' productivity. Consistent with prior theory, we find a positive effect of transparent performance data on productivity, but it is driven by employees avoiding the least productive behaviors rather than increasing the most productive ones. We also investigate two novel relational moderators of the relationship between transparent performance data and productivity—supervisor support (vertical relation) and social comparison orientation (horizontal relation)—which we believe are understudied, given their increased importance in more transparent performance environments. We find that both moderate the positive effect: employees with low supervisor support and low social comparison orientation benefit most from transparent performance data, suggesting that it acts as a substitute for both managers and informal social comparison. We discuss the implications for theory and for the practice of designing transparent feedback systems.

KEYWORDS: performance feedback, transparency, productivity, field experiment

Organizations have long used performance feedback—information about the effectiveness of one's work behavior (Ashford & Cummings, 1983; Taylor, Fisher, & Ilgen, 1984)—to improve employee productivity (Alvero, Bucklin, & Austin, 2001; Balcazar, Hopkins, & Suarez, 1985; Ilgen, Fisher, & Taylor, 1979; Larson, 1989). Managers use traditional feedback interventions, such as an annual performance review, to provide knowledge of results (Ammons, 1956) and to "cue" efforts to improve them (Vroom, 1964).

Yet what is deemed performance feedback is shifting. Performance data that is deemed "transparent"—*more detailed, more real-time, and shared with a broader audience of employees*—is increasingly expected to do the job of traditional one-on-one manager-led performance reviews. Today, handheld computers (Amazon) or wearable devices (Tesco) track and optimize employees' every move (Head, 2014; Kantor & Streitfeld, 2015; Rawlinson, 2013), point-of-sale systems scrape all transactions for signs of employee fraud (Pierce, Snow, & McAfee, 2015), hundreds of embedded sensors in UPS trucks record the truck's and its driver's every action to unearth and enforce time-saving tactics (Goldstein, 2014), and gamified leaderboards make such tracked metrics public in real time (Bernstein & Blunden, 2015; Mollick & Rothbard, 2014). Studies suggest such transparent performance data has already displaced traditional performance feedback in nearly 10 percent of *Fortune* 500 companies (Cunningham & McGregor, 2015).

While it is assumed that such transparency will raise productivity (e.g., Huhman, 2014; Lublin, 2011), theory and empirical evidence have yet to catch up with practice. This uncritical enthusiasm is not a new problem for the field of performance feedback; there was, for example, a "persistent and unwarranted belief that feedback intervention always improves performance" (1996: 277) until Kluger and DeNisi's seminal review demonstrating "a considerable body of evidence" that the effect of traditional performance feedback interventions on performance is "quite variable" (1996:

254) and that, in fact, “in over one third of the cases feedback interventions reduced performance” (1996: 275). The authors highlighted, through the conceptual lens of their Feedback Intervention Theory (FIT), the wealth of moderators—both attributes of the feedback (e.g., framing, specificity, frequency) and complementary features attached to the feedback (e.g., goal setting, rules, incentives)—that can make traditional, manager-led performance feedback more effective. As traditional conceptions of feedback intervention give way to new technology for making performance data transparent, it is time for a similar scholarly effort, focused on what is new.

In this article, we introduce two conceptual distinctions with which to begin bridging the traditional performance feedback literature to the emerging world of transparent performance data. First, we distinguish between the impact of transparent performance data on *productive* versus *nonproductive behaviors*. Traditionally, managers choose how to deliver feedback and prior research has highlighted the importance of positive versus negative framings or cues. But in a world of transparent performance data, a manager’s choice of which behaviors to track becomes more important, so we extend beyond framing to distinguish between the tracked behaviors themselves. Second, we introduce *relational moderators*—moderators based on an employee’s relations to coworkers (such as his or her boss, peers, and subordinates)—as a key new type of moderator of the effectiveness of transparent performance data. Because traditional performance feedback interventions typically involve a defined set of individuals (most often, an employee-supervisor dyad), prior research has focused on cognitive or emotional moderators of effectiveness (for a comprehensive historical review, see Kluger & DeNisi, 1996). The multiplex nature of transparent performance data—employee-all rather than just employee-supervisor—broadens and deepens the importance of a subset of social moderators, which we call relational moderators.

We draw on a field experiment, including embedded participant observation, involving a large gas utility in the southeastern United States. A randomly selected group of frontline employees went from traditional supervisor-employee performance feedback to also having direct, transparent access to individual performance data (their own and their peers'), automatically tracked by devices and by their trucks. Because this workforce was not subject to strong financial incentives for performance but did work sufficiently complex to be representative of any job requiring professional capabilities and judgment, we could separately instrument and analyze the direct effect of the switch from traditional performance feedback to transparent performance data.

Our research makes three theoretical contributions. First, by providing field-based, controlled, empirical evidence on the effect of transparent performance data on employee productivity relative to traditional performance feedback, we depart from the literature's traditional focus on manager-led feedback interventions and also address the call for controlled field studies (Kluger & DeNisi, 1996). The feedback interventions literature, much of it based on lab studies, has paid scant attention to how individuals might respond to transparency-based (data-based) rather than manager-based (conversation-based) interventions in the field, even as the latter is becoming popular in practice (Ewenstein, Hancock, and Komm, 2016).[1] This emphasis on transparency opens up new lines of research on performance feedback and productivity.

Second, we advance the current understanding of how transparent performance data might have differential impacts on productive and nonproductive behaviors. Conventional wisdom would suggest that transparent performance data can improve performance by increasing attention to task motivation (by providing benchmarks and goals), task learning (by identifying the high-

[1] As an illustration, the most popular *McKinsey Quarterly* article in 2016 was about organizations which have moved from traditional performance feedback to transparent performance data.

performance coworkers from whom one can learn), or self-development (by activating meta-task processes) (Kluger & DeNisi, 1996). Yet the pride felt when others see our excellence and the shame felt when others see our poor performance are different phenomena (Lewis, 1971) and no research has looked at differential effects of transparency on productive and nonproductive behaviors. Our study cleanly distinguishes the two and, in finding differential impact, contributes to theories on the motivation, learning, and attentional aspects of performance feedback.

Third, this research, by highlighting an important new type of moderator—the relational moderator—adds a social component to the cognitive and emotional moderators previously studied. When an employee's performance is made transparent, the effect on that employee is likely to be driven not just by the feedback itself but also by his or her relations with supervisors and peers—in this case, the employee's level of supervisor support (vertical relation) and social comparison orientation (horizontal relation). The identification of these moderators can help scholars and managers understand and mitigate some of the variance that has been common in transparent performance feedback interventions. We elaborate on these contributions in the discussion.

THEORETICAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

Performance Feedback: Theory and Evidence

Numerous studies in psychology, organizational behavior, and operations management have demonstrated the potentially beneficial effects of feedback on individual performance in organizations (e.g., Alvero et al., 2001; Balcazar et al., 1985; Dockstader, Nebeker, & Shumate, 1977; Fedor, 1991; Florin-Thuma & Boudreau, 1987; Guzzo, Jette, & Katzell, 1985; Ilgen et al., 1979; Ivancevich & McMahon, 1982; Pritchard, Bigby, Beiting, Coverdale, & Morgan, 1981; Taylor et al., 1984). For example, Ilgen et al. (1979: 349) stated in their review that feedback is "essential for learning and for motivation." The numerous mechanisms by which feedback triggers learning and motivation (and therefore improves performance) include the law of effect (reinforcement and punishment) (Thorndike, 1927), control theory (Annett, 1969), goal setting theory (Locke & Latham, 1990; Locke, Shaw, Saari, & Latham, 1981), the multiple-cue probability learning paradigm (Balzer, Doherty, & O'Connor, 1989), social cognition theory (Bandura, 1991), and attention (Kluger & DeNisi, 1996), among others.

However, a long tradition of empirical research also demonstrates that performance feedback systems can undermine performance, eroding it for the best employees (Haas & Hayes, 2006), the worst employees (Podsakoff & Farh, 1989), or anyone in between (Ivancevich & McMahon, 1982; Kluger, Lewinsohn, & Aiello, 1994). Most notably, Kluger and DeNisi's (1996) meta-analysis of empirical studies, accounting for 607 effect sizes and 23,663 observations, found that feedback interventions improved performance on average but that 38 percent of them decreased performance. Those mixed results prompt Kluger and DeNisi (1996) to generate a Feedback Intervention Theory (FIT) and test their propositions against a long list of moderators using a meta-analysis of past studies. Despite a hodgepodge of studies not designed for the questions that Kluger and DeNisi were asking, they still find significant—although weakly significant—confirmation that cues directing attention to meta-task processes tend to attenuate the effect of feedback, while those directing attention to task-motivation or task-learning processes augment its effect:

> 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 (velocity FI and goal setting), and is 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].

We believe FIT offers a comprehensive framework to evaluate a century's worth of empirical evidence of how traditional performance feedback interventions affect performance. Our goal in this study is to extend FIT by (a) applying it to the new questions raised by transparent performance data and (b) empirically proposing behavioral and relational factors that would make the theory more complete for the evolving nature of performance management.

New technology, new questions. Workplace transparency is becoming widespread (Bernstein, 2014, 2017). Advancements in tracking technologies have permitted organizations to collect increasingly detailed, real-time performance data, which are traditionally provided to managers to summarize and deliver to employees in periodic performance feedback discussions. But the same advancements in technology have also permitted those performance data to be delivered directly to employees (rather than through a conversation with a manager), allowing more people to see more information, more frequently, about their own performance and the performance of others.

The resulting transparent performance data, relative to traditional, manager-led performance feedback, embody many of the positive features summarized in Kluger and DeNisi's conclusion quoted above. They are typically (a) detailed and specific, with multiple cues to support learning; (b) blunt and without emotion (system-generated), clearly displaying the task-level difference between one's performance and established standards and thus less likely to become personal (and potentially more likely to be viewed as objective); and (c) capable of extremely high velocity (frequency). Perhaps the most attractive feature of transparent performance data, however, is their goal-setting function. In prior research on traditional performance feedback (including the FIT meta-analysis), one of the most consistent significant findings is the importance of pairing feedback interventions with goal setting (e.g., Balcazar et al., 1985; Becker, 1978; Ivancevich & McMahon, 1982; Pritchard et al., 1981). Yet while this improves performance on single-criterion tasks studied in the laboratory (e.g., Becker, 1978; Komaki, Barwick, & Scott, 1978; Latham & Locke, 1991; see also Locke et al., 1981 for a thorough review), such single-criterion tasks rarely exist in modern organizations. That makes traditional approaches to pairing goal setting and performance feedback impractical or, worse, converts goal setting into the time-consuming yet vapid check-the-box game of impression management hidden within most annual performance reviews (Culbert & Rout, 2010). Making performance data transparent is therefore becoming a substitute for goal setting and, increasingly, for traditional manager-led performance feedback: organizations are replacing the negotiation of particular performance criteria with carefully instrumented, tracked, and displayed activity that each employee can access (Buckingham & Goodall, 2015), literally seeing where he or she stands within the organization.

Given the capacity of transparent performance data to achieve so many positive attributes found in prior research, we hypothesize:

Hypothesis 1 (H1): Substituting transparent performance data for traditional performance feedback (manager-led performance reviews) increases overall employee performance.

***New technology, new distinctions*.** Research has investigated how positive or negative framing (McFarland & Miller, 1994) of traditional performance feedback (for example, telling an employee at the 15th percentile either that he or she is in the top 20 percent or that he or she is *not* in the top 10 percent) may moderate the relationship between feedback and performance (e.g., Ashford & Tsui, 1991). Positive framing, for example, may increase intrinsic motivation by enhancing feelings of competence (Sansone, 1986), but may also reduce pressure and arousal (Kluger et al., 1994), thus lowering performance (Haas & Hayes, 2006). Transparent performance data, because it is more data than framing, refocuses the decision from how to frame the conversation positively or negatively to which behaviors (productive or nonproductive) to track and report. The question is whether the type of behavior made transparent will similarly moderate the degree to which doing so impacts performance.

For productive behaviors (those that positively contribute to performance, such as working more efficiently), at least in environments in which individuals must allocate attention (Kluger & DeNisi, 1996) or effort (Hannan, McPhee, Newman, & Tafkov, 2012) across multiple tasks, making performance transparent might improve motivation but simultaneously encourage unproductive distortions (e.g., gaming) in how work is done, such as failing to increase performance except when close to a bonus threshold (Delfgaauw, Dur, Non, & Verbeke, 2014); playing it safe at the average and avoiding risks, due to strong negative incentives (like being fired) (Brown, Fisher, Sooy, & Sprinkle, 2014); or even sabotaging others' work or artificially inflating one's own short-term performance (Charness, Masclet, & Villeval, 2013).

For nonproductive behaviors (those that fail to positively contribute to performance, such as working less efficiently or idle time), there is scant evidence, but studies on counterproductive behaviors (those that directly undermine performance) may be indicative. For example, using multi-firm field experiment data from tens of thousands of waiters in hundreds of restaurants, Pierce et al. (2015) found that using tracking technology to detect theft not only reduced it by $24 per week but also increased total revenue by $2,975 per week. Staats, Dai, Hofmann, and Milkman (2016) found that RFID-based electronic monitoring of handwashing in hospitals had an initial positive effect on compliance ranging from 20 to 60 percent, although compliance then declined over time and even fell below pre-intervention levels after monitoring was removed. These studies show that making counterproductive behaviors (e.g., theft, noncompliance) transparent significantly reduces them since people do not want to be caught or shamed (Lewis, 1971; Taylor, 1985), which is likely to be true for nonproductive behaviors as well.

We interpret the research above to suggest that transparency-based feedback may be most powerful when applied to nonproductive behaviors. We believe that difference is likely accentuated in a *transition* from traditional performance feedback to transparent performance data, because nonproductive behaviors are less frequently discussed in traditional performance feedback systems—managers systematically shy away from sharing nonproductive feedback (Larson, 1986). Therefore, just as previous research has found that the framing moderates feedback's effectiveness in increasing performance (Higgins, 1987; Van-Dijk & Kluger, 2004), with negative framing often but not always more powerful (Kluger & DeNisi, 1996), we hypothesize that which behaviors technology makes transparent may impact how well it increases performance:

Hypothesis 2 (H2): Substituting transparent performance data for traditional performance feedback triggers a greater behavioral shift towards productivity for nonproductive behaviors than for productive behaviors.

New technology, new moderators. Making performance transparent tends to involve making the traditional performance feedback conversation far more public. Transparent performance data, by definition, implicate a wider spectrum of observers than a traditional performance feedback review does. That suggests a new set of potentially important *relational moderators*, which capture the relationship between the individual and the public. In this study, we investigate two relational moderators—supervisor support and social comparison orientation—that are well established in the literature on organizations but, so far as we are aware, have not been studied in the contexts of performance feedback or transparent performance.

Supervisor Support: Vertical Relationships Moderate the Effects of Transparent Performance Data

One key relational difference between traditional performance feedback and making performance transparent is the role of the supervisor. Traditionally, a supervisor is the employee's primary source of performance feedback, even if that feedback has been gathered from others, as in a 360-degree review (DeNisi & Kluger, 2000). In short, the supervisor delivers the feedback, filtered and framed as deemed appropriate by the supervisor. With transparent performance, however, the data goes straight to the employee, bypassing the supervisor's filter. How will that affect performance? Conceptually, it depends on whether transparent performance data is a *complement to* or a *substitute for* the supervisor's role.

On the one hand, transparent performance data can only reveal where one stands, not how to improve. Although transparent performance data have been framed as a replacement for traditional performance reviews, a performance review is more than just data—a supervisor can provide multiple forms of guidance in such discussions. Transparent performance data may therefore complement coaching and sensemaking from a supervisor that translates data from an opportunity for improvement into concrete steps for learning and performance (Pfeffer, 2016). Counterproductively, providing transparent performance data without supervisor support could trigger feelings of helplessness (Martinko & Gardner, 1982) and thus undermine performance.

On the other hand, transparent performance data may substitute for the supervisor's role (Hamel, 2011; Tapscott & Ticoll, 2003). Finding out where you stand in comparison to coworkers can provide you with multiple role models—high-performing peers—who may even be better coaches than your supervisor (e.g., Ilgen, Peterson, Martin, & Boeschen, 1981; Larson, 1986). Like transactive memory (Ren & Argote, 2011; Wegner, 1987), seeing where you stand could be a way to identify where the expertise lies (Song, Tucker, Murrell, & Vinson, 2016).

Perceived supervisor support—that is, an employee's view on how much his or her supervisor values his or her contributions and development (Kottke & Sharafinski, 1988)—offers a validated, meaningful instrument through which to test this relationship between transparent performance data and future performance. Perceived supervisor support has been empirically tied to in-role and extra-role performance (Eisenberger, Stinglhamber, Vandenberghe, Sucharski, & Rhoades, 2002;

Jokisaari & Nurmi, 2009; Shanock & Eisenberger, 2006), but the effect of transparent performance on that relationship remains unstudied. If transparent performance data complements the supervisor's traditional role, then the performance of employees with high perceived supervisor support—employees whose supervisors are already adeptly supporting their improvement—should improve more. On the other hand, if transparent performance data substitutes for the supervisor's traditional role, then the performance of employees with low perceived supervisor support—employees whose supervisors are not already adeptly supporting their improvement—should improve more. Our reading of the literature suggests that supervisor support has as much potential to moderate the relationship between transparent performance and productivity positively as negatively. We therefore offer paired hypotheses:

Hypothesis 3a (H3a): Substituting transparent performance data for traditional performance feedback complements supervisors, so employees with more-supportive supervisors improve their productivity more than those with less-supportive supervisors.

Hypothesis 3b (H3b): Substituting transparent performance data for traditional performance feedback substitutes for supervisors, so employees with less-supportive supervisors improve their productivity more than those with more-supportive supervisors.

Social Comparison Orientation: Horizontal Relationships Moderate the Effects of Transparent Performance Data

A second key relational difference involves the relationship between an employee and his or her peers. Festinger's (1954) seminal paper proposed that, in the absence of clear standards, people evaluate themselves, their opinions, and their capabilities in comparison with similar others (see Gibbons & Buunk, 1999; Suls, Martin, & Wheeler, 2002; and Wood, 1989 for detailed reviews of social comparison orientation theory). Although the "desire to learn about the self through comparison with others is universal" (Gibbons & Buunk, 1999: 199), the extent to which people do so varies with the established scale of social comparison orientation, the Iowa-Netherlands Comparison Orientation Measure (INCOM) (Gibbons & Buunk, 1999). While social comparison orientation has been linked to the productivity effect of performance transparency in the lab (McFarland & Miller, 1994), we are aware of no field experiments on social comparison orientation. Yet a field experiment has the advantage of involving employees who have worked together for much longer than a brief lab experiment.

If the "primary goal of social comparison is to acquire information about the self" (Gibbons & Buunk, 1999: 129) and if such information is typically used to evaluate and improve oneself, it would seem to follow that those more prone to social comparison are more likely to increase their productivity when given more performance transparency. Indeed, in a recent study of call center employees, Lount, Jr. & Wilk (2014) show that social comparison, triggered by posting individual performance, increased productivity in groups. We therefore hypothesize:

Hypothesis 4a (H4a): When substituting transparent performance data for traditional performance feedback, employees with a stronger social comparison orientation improve their productivity more.

However, in real work environments (as opposed to lab contexts), it's quite possible that, even before employees formally receive any feedback, they already have a sense of how they compare with each other—from their supervisors' comments, from feedback-seeking behaviors (Ashford, 1986; Ashford, Blatt, & Walle, 2003; Ashford & Tsui, 1991) such as informal discussions with supervisors or peers, or even just from observing each other.

To the extent that a person with a greater social comparison orientation would be more likely to have already gathered performance information about others before it was provided *formally*, the effect of formally providing transparency could be smaller. It is also possible that heightened social comparison, combined with increased performance transparency, could lead to negative psychological consequences, such as diminished trust in coworkers (Dunn, Ruedy, & Schweitzer,

2012) and discouragement (Beshears, Choi, Laibson, Madrian, & Milkman, 2015), and to unproductive behaviors. We therefore hypothesize:

> *Hypothesis 4b (H4b): When substituting transparent performance data for traditional performance feedback, employees with a weaker social comparison orientation improve their productivity more.*

METHODS

Research Setting

The context of our study is a service operation—a natural gas distribution company (referred to as GasCo, a pseudonym) with 1,100 employees serving approximately 425,000 residential, commercial, and industrial customers in the southeastern United States. Most of GasCo's 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 requests from customers to turn on gas, turn off gas, repair gas appliances, repair leaks, and respond to emergencies. They typically start their day by logging into the system on their trucks and accepting orders. A mechanic maps a path to an order (for example, a customer's house), arrives on site, completes the order, then drives to the next order or takes a break. Although the tasks may appear routine, mechanics identify themselves as professionals because of (a) the risk inherent in any activity involving gas, (b) their substantial training, and (c) the wide variability in the contexts, systems, and devices they are expected to safely diagnose and fix. Their tasks therefore fit our two task-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, the mechanics interact with customers 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." Second, and equally important, GasCo's work force 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 than they are in the real workplace (Holmstrom & Milgrom, 1991; Larkin, Pierce, & Gino, 2012). Although many, if not most, front-line jobs lack strong financial incentives, such incentives are part of every other field experiment on performance transparency of which we are aware. GasCo's employee incentive plan, based on companywide objectives, was thus relatively disconnected from individual performance. This allowed us to observe the effect of performance transparency *itself*, decoupled from financial incentives or fear of career consequences.

Like most US gas utilities, GasCo had grown by acquiring dozens of municipal gas utilities. Not long before our study, it had consolidated the customer service organizations (including the mechanics) of all its acquisitions. GasCo's mechanics, all of whom were represented by one of two unionized bargaining units, 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 generated a single scorecard of metrics—collected automatically by technology in the mechanics' trucks and computers—to which mechanics had collectively agreed. The metrics included three categories of mechanic-specific metrics on efficient allocation of time (percentages of productive time, support time, and nonproductive time, defined in the "Measures" subsection).

Data

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—a GasCo intranet page with a scorecard displaying transparent performance data visible to all workers in one's own work center. This is the same data that supervisors previously received (and continued to receive) and would, at their discretion, discuss with mechanics. The other seven work centers served as a control group operating at the status quo: one received feedback from one's supervisor, who continued to have access to all of the data which became, in the treatment condition, transparent to all mechanics in the work center. 31 mechanics were thus randomly assigned to the treatment condition and 92 to the control group. Thus the only difference between the two conditions was who received the data: just the supervisors, or the supervisors and the mechanics within that work center.

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

Figure 1 shows a sample screenshot of the daily scorecard information visible to a mechanic in the treatment group (here, with the employees' names disguised). Employee 1, for example, could see not only his or her own performance metrics for the previous day, but also the metrics for everyone else in his or her work center. Mechanics in the treatment group received an email every morning with a link to the scorecard information, which they could access from their truck or from any computer. We tracked how often the intranet webpages were accessed to ensure the quality of the intervention—that is, to be sure the mechanics were actually accessing the performance data—although, due to both technological and human subjects limitations, we could not identify who had made any particular visit to the data. The experimental pilot ran from June

25 to August 29, 2014. We retrieved daily performance data from GasCo's archive for June 1, 2013 through August 29, 2014; that is, 389 calendar days before the intervention and 65 working days during the pilot.

Survey. In order to measure perceived supervisor support and social comparison orientation, we administered a pre-experimental survey 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. Consistent with many studies in organizational behavior, we used INCOM (Gibbons & Buunk, 1999) to measure social comparison orientation.

As an additional control, we asked the mechanics to assess their own past performance on a variety of metrics, given prior evidence that one's perception of one's own past performance influences one's expectations for future performance, which, in turn, have been shown to moderate the productivity impact of performance feedback (Northcraft & Ashford, 1990). For example, we asked, "[C]ompared to all 123 mechanics at GasCo, I think my performance would rank me ___ out of 123." By pairing a mechanic's response with actual past performance before the intervention, we could control for past performance, both actual and perceived.

We sent the survey by email on June 19 (six days before starting the experimental intervention) and gave the mechanics four days to complete it. They could access it from their truck laptops or from any computer. We made clear that the survey was conducted by the researchers, not by 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 123 mechanics and received 63 responses, a response rate of 47 percent, which management reported was typical for this population. Comparing human-resources data for responding and nonresponding mechanics revealed no bias in the type of individual who responded.

We timed the survey to be a trigger, in both the treatment and control groups, for any Hawthorne effects (Roethlisberger & Dickson, 1939), whereby performance is changed by the mere fact of attention or environmental modification rather than by the treatment itself. By simultaneously triggering any such effects in both the treatment and control groups, we alleviated concerns that results in one condition relative to the other might be due to a Hawthorne effect.

Participant observation. In the spirit of the longstanding qualitative tradition of participant observation (Becker, 1958; Jorgensen, 1989; Spradley, 1980), one researcher was embedded into the workforce for the week of June 30, the second week of the field experiment. He was selected for his ability to fit in with traditional recruits for the service mechanic role at GasCo, but he was trained in how to properly collect field notes (Emerson, Fretz, & Shaw, 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 mechanics in different work centers. His note-taking was not seen as unusual, but rather as typical for an apprentice. In the evenings, he synthesized his notes and added detail while memories remained fresh. Because mechanics spend so much time driving from site to site, 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 survey.[2]

Measures

Table 1 shows the definitions and descriptive statistics for our key variables.

-------------------- Insert Table 1 About Here --------------------

[2] A number of anonymized quotations from mechanics are incorporated into the results and discussion sections of this paper, all of which were collected via participant-observation.

Dependent variables. To measure performance, we adopted GasCo's 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 work day. Both productive time and support time activities were clearly and stringently defined at GasCo (Figure 1), with nonproductive time as the default residual; if an activity was not defined as productive or support, it was classified as nonproductive.

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 and logic: en route time was constrained based on route and traffic data, while onsite time was constrained based on standardized times for the requested services. That is, a trip from here to there counted as 20 minutes of productive time if that's how long the route and traffic data indicated such a trip should take, regardless of how long it actually took on any given day. Support time—maintenance, training, preparation, and colleague support—was similarly constrained. Permissible support time could be bundled with certain productive time activities (such as picking up materials and loading or unloading a truck), allocated after a certain number of total hours had been worked (for example, on training or on union business), assigned based on work schedules (for example, a meeting), or triggered by management request (for example, building maintenance) or by circumstances (for example, vehicle maintenance).[3] While mechanics understood how the metrics were calculated, they did not directly allocate time to one category or another, but instead engaged in activities that were automatically tracked and coded for one of those three categories. To prevent abuse or gaming,

[3] An alternative way to view these metrics is that *% Productive Time* represents time spent to increase current productivity, *% Support Time* represents time spent on activities, such as maintenance and training, that might increase future productivity or work quality, and *% Nonproductive Time* represents time spent on activities that are unlikely to increase productivity or work quality.

time allocation and activity records were audited. Indeed, one reason GasCo thought transparency might be productive was that it might reduce the possibility of abuse by increasing the number of eyeballs on the data.

GasCo designed these metrics to help mechanics allocate time productively. Our decision to use time allocation as a measure of productivity was therefore 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) to make wise use of discretion in allocating their time. But we also found support for time allocation as a productivity measure in several influential literatures, including operations management scholars studying scheduling (Pinedo, 2012) in factories (Berman, Larson, & Pinker, 1997), healthcare (Kc & Terwiesch, 2009), trucking (Roberti, Bartolini, & Mingozzi, 2014), and financial services (Staats & Gino, 2012). All three metrics—*% Productive Time*, *% Support Time*, and *% Nonproductive Time*—were in use at GasCo 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.

In our empirical analysis, *% Nonproductive Time* and *% Productive Time* are the dependent variables. Because a work day is allocated entirely between the three categories, showing changes in two is sufficient to reflect any change in how mechanics allocate their time. *% Nonproductive Time* is a negative productivity metric; changes represent mechanics' decisions to spend more or less time on useful activities to increase current or future productivity. Once mechanics decide to spend time in useful activities, they allocate it between *Productive Time* and *Support Time* and we analyze *% Productive Time* to investigate their allocation across the two. Throughout the period of our study, there was ample excess customer and GasCo demand for both productive and support activities, leaving mechanics sufficient room to improve any of their metrics.

Moderating and Control Variables. Our measures for perceived supervisor support and social comparison orientation, consistent with their design and previous use, are calculated as the sum of the scores from each survey question on supervisor support or social comparison orientation, adjusted for reverse coding. We also incorporate control variables based on demographic information from GasCo's internal records, including tenure, age, gender, and race.

We control for past performance in all regressions. As an example of how past performance can affect the relationship between productivity and the transparency of performance data, consider findings that feedback can positively influence poorer performers while having little influence on better performers (Pritchard et al., 1981), meaning that those who do not receive positive recognition are primarily responsible for an increase in organizational performance, through peer pressure and a wish to conform with higher-performing peers (e.g., Bradler, Dur, Neckermann, & Non, 2013; Schultz, Juran, & Boudreau, 1999). Indeed, people with different past performance respond differently to different performance feedback mechanisms, including social comparison (McFarland & Miller, 1994). Because some of these effects could be triggered either by actual past performance or by the gap between perceived and actual past performance, we incorporate both as controls. Our measure of a mechanic's actual past performance comes from archival data on *% Nonproductive Time* and *% Productive Time*; self-evaluated past performance is a standardized self-evaluation, between 0 and 100, based on the survey.

Analytical Strategy and Main Regression Model

To analyze the field experiment data, we used a difference-in-differences estimation model, which allows a precise yet simple analysis of the effect of the intervention on the treatment group relative to any changes experienced by the control group during the same period. The model is:

$$Y_{it} = \alpha + (\beta_1 \times \text{Treatment}_{it}) + (\beta_2 \text{ x } \text{Post}_{it}) + (\beta_3 \text{ x } (\text{Treatment}_{it} \times \text{Post}_{it})) + \sum Controls + \varepsilon_{it}. \quad (1)$$

Y_{it} is the dependent variable: the performance metric at the employee(i)-workday(t) level. Treatment_{it} is an indicator variable that equals 1 if the employee was working in one of the four pilot sites. Post_{it} is an indicator variable that equals 1 if the date was on or after June 25, 2014, when the treatment began. The main estimation uses ordinary least squares (OLS) regressions. Consistent with Bertrand, Duflo, and Mullainathan (2004), standard errors of the coefficients are computed using the block-bootstrap method (clustering by work center). If the effect of the treatment 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*.

RESULTS

Descriptive Statistics

On average, mechanics' nonproductive (productive) time over the sample period is 30.70 percent (59.13 percent) with good variation (standard deviation is 23.41 percent for nonproductive time and 24.36 percent for productive time). The median age and median tenure are 50 and 19 years, respectively.[4] The survey data showed good variation in the measures based on the ratio of standard deviation to the mean. The only data 'anomaly' is that the mean value of the self-evaluation is 67 percent rather than 50 percent, consistent with the well-documented tendency to overestimate one's own relative performance (Kruger & Dunning, 1999).

[4] Comparing all pre-experiment variables across the treatment and control groups, our randomization appears to have been successful, with the exception of two variables for which the ratio of the difference between the treatment and control groups' means was greater than 10 percent. We explicitly control for these unbalanced variables in our statistical analysis.

-------------------- Insert Table 2 About Here --------------------

Table 2 tabulates the unconditional correlations between the dependent and independent variables. *% Nonproductive Time* is positively correlated with age, past performance on *% Nonproductive Time*, and being white and is negatively correlated with past performance on *% Productive Time*, self-evaluation, and perceived level of supervisor support. *% Productive Time* is negatively correlated with tenure, age, past performance on *% Nonproductive Time*, and being white and is positively correlated with social comparison orientation, past performance on *% Productive Time*, self-evaluation, and perceived level of supervisor support. Tenure is negatively correlated with social comparison orientation, self-evaluation, and perceived level of supervisor support. All these correlations are statistically significant at the one-percent level.

Baseline Result (H1 and H2): Does Performance Transparency Improve Productivity?

-------------------- Insert Table 3 and Figure 2 About Here --------------------

Table 3 shows the main regression results. Columns 1, 3, and 5 use *% Nonproductive Time* as the dependent variable, while Columns 2, 4, and 6 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, 3, and 5 indicate a negative effect on *% Nonproductive Time*; that is, mechanics in the treatment group, relative to those in the control group, chose to spend more of their time on useful activities. Compared with the pre-intervention mean value, the results in Column 3 indicate an 11.01-percent decrease in *% Nonproductive Time,* suggesting that substituting transparent performance data for traditional performance feedback improved productivity at GasCo (H1).

Columns 2, 4, and 6, however, present an important caveat. The effect of the intervention on *% Productive Time* was slightly positive (a 4.77-percent increase in Column 4, considering the pre-intervention mean value), but not statistically significant after we block-bootstrapped the standard errors, clustered by work group. While mechanics in the treatment condition chose to spend more of their time on useful activities, more of that time went to support activities than GasCo had hoped or expected. We visualize these results in Figure 2.

Comparing Columns 1, 3, and 5 with 2, 4, and 6, the treatment effect on nonproductive time is greater than that on productive time, in both magnitude and significance; mechanics focus on achieving less nonproductive time, not on achieving more productive time This is consistent with H2: transparent performance data on nonproductive behaviors triggers a greater behavioral shift towards productivity than similar performance data on productive behaviors.

To ensure the quality of our intervention—that is, to be sure that access to the performance data and not something unrelated is driving the results—we obtained the number of views of the scorecards for each of our four treatment sites during the experiment period. Across the four sites, the average employee accessed the daily report at least once every three days and in some cases far more frequently. The sites at which mechanics accessed the report most often were also the ones with the greatest productivity boost, indicating that it was, indeed, access to the transparent performance data that drove the treatment effects.

Transparent Performance Data Substitutes for Supervisor Support (H3)

To test H3a and H3b, we ran the baseline regressions, splitting the sample (Baron and Kenny, 1986; Jaccard and Turrisi, 2003) into those who reported a level of higher or lower perceived supervisor support (either greater than or equal to the sample median or lower than the median);

these are labelled "High Supervisor Support" and "Low Supervisor Support," respectively, in Table 4. The objective in this analysis is to examine whether the subsample of mechanics with supervisors characterized as supportive benefitted more or less from the transition to transparent performance data than the subsample with supervisors characterized as unsupportive. We chose to split at the median because, as the mechanics told our participant observer, "there are great bosses and bad bosses… and they split about fifty-fifty."

We find treatment effects of greater magnitude and statistical significance for both *% Nonproductive Time* and *% Productive Time* for those who perceived a low level of supervisor support, consistent with H3b instead of H3a and suggesting that transparent performance data may serve as a substitute for supervisors. Compared with the pre-intervention mean values for both productivity metrics, these results indicate an 18.71-percent decrease in *% Nonproductive Time* and a 16.43-percent increase in *% Productive Time*. Our participant-observer heard one mechanic say that he liked it that "low performers could approach high performers"—knowing now who they were—to learn how to improve if a mechanic "didn't have a good supervisor."

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

Transparent Performance Data Substitutes for Social Comparison (H4)

To test H4a and H4b, we ran the baseline regressions, similarly splitting the sample into those who reported a level of higher or lower social comparison orientation (either greater than or equal to the sample median or lower than the median); these are labeled "High Social Comparison" (Columns 1 and 3) and "Low Social Comparison" (Columns 2 and 4) in Table 5. The objective in this analysis is to examine whether the subsample of higher-social-comparison-orientation mechanics benefitted more or less from the transition to transparent performance than the subsample of lower-social-comparison-orientation mechanics. As with previous studies using social comparison orientation (e.g., Michinov & Michinov, 2001; van Quaquebeke, van Knippenberg, & Eckloff, 2011), the median provides the most sensible split between high and low.

The negative treatment effect on *% Nonproductive Time* is greater—both in magnitude (a 17.36-percent reduction from the pre-intervention mean) and in statistical significance—for the low-social-comparison group. As one supervisor commented to the participant-observer, "I had always told my workers they could come to me to see [the data]," so those who did—the higher social comparison orientation individuals—had less to gain. We find support for H4b instead of H4a.

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

Robustness Checks

We ran robustness checks to address three characteristics of our field experiment.

First, the gas utility business is seasonal. The effect of seasonality, in part, can be seen through the significant, negative coefficient on our time variable, *Post*, in a number of 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 treatment effects were similar.

Second, our dependent variables, *% Nonproductive Time* and *% Productive Time*, are correlated (though not perfectly, since there is also the third category, "Support Time"), which means that 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 block-bootstrapped standard errors (clustered by work location).

DISCUSSION

> Sixty years ago, digital computers made information readable. Twenty years ago, the Internet made it reachable. Ten years ago, the first search engine crawlers made it a single database. Now Google and like-minded companies are sifting through the most measured age in history, treating this massive corpus as a laboratory of the human condition. They are the children of the Petabyte Age. The Petabyte Age is different because more is different (Anderson, 2008: 1).

Performance transparency once meant the disclosure of organizational outcomes, but the steady advance of enabling technologies (for a review, see Kidwell & Sprague, 2009) has made it increasingly possible to make an individual employee's activities transparent. Such transparency is increasingly open to all, not just managers. While the ability to track more information does not necessitate open access to it, the two seem to have evolved together. That has implications for many aspects of management, as who gets the data can determine how data are used, but the effects are particularly acute for the field of performance feedback, where employee-supervisor discussions are being displaced by employee performance transparency.

Our goal in this study is to examine whether replacing traditional manager-led performance feedback with transparent performance data (system-generated, high frequency, and visible to a larger audience) affects employee performance. In a field experiment in a large US service organization, our main results show that transparent performance data improves productivity, mostly by decreasing the amount of time employees spend on nonproductive behaviors. Our findings further indicate that two relational variables—that is, variables based on the relationship between the employee and others—moderate the effect of transparent performance data on employee productivity. In both cases, transparent performance data operated as a substitute for human relationships: employees with lower perceived supervisor support (the vertical moderator) and employees with lower social comparison orientation (the horizontal moderator) had significantly greater productivity gains when traditional performance feedback was replaced with transparent performance data.

Theoretical Implications

Our results advance the literature on performance feedback. First, we directly link the literature on traditional performance feedback with the emerging phenomenon of transparent performance data, examining whether replacing the former with the latter affects employee performance. Our findings support the notion that, on average, existing theory on traditional performance feedback applies to the new world of transparent performance data: as existing literature would predict (H1), seeing coworkers' performance data in real time improves performance. Yet there are important nuances. Transparent performance data did not encourage an increase in the "best" behaviors (as perceived by the mechanics and captured by the participant-observer) as much as it encouraged an avoidance of the "worst." Put differently, it made the mechanics in our study more determined *not* to stand out for their nonproductive time, but not more determined to stand out for their productive time. In their own words, the transparent performance encouraged mechanics to "hide in the middle of the pack" and to "conform, not excel." The social nature of transparent performance data appears to drive employees to focus their attention on the self (Kluger & DeNisi, 1996), which triggers motivations for compliance and conformity (Cialdini & Trost, 1998) rather than for excellence.

Second, we introduce and demonstrate the increasing importance of relational moderators—moderators based on an employee's relations to others at work (such as his or her boss, peers, and subordinates)—as workplaces become more transparent. Traditional performance feedback is dyadic (employee-supervisor). Transparent performance feedback is multiplex (employee-all), reflected in the significance of our vertical and horizontal relational moderators. These results show that transparent performance data, like transparent goals, can interact with social situations in a variety of ways (Endler, 1993). It is therefore increasingly important that research on performance feedback incorporate relational considerations (Grant & Parker, 2009).

Our relational moderator for vertical relations, supervisor support, showed a significant inverse relationship with productivity improvement for *both productive and nonproductive time* in the transition from traditional performance feedback to performance transparency. In other words, supervisor support and performance transparency were substitutes (Tapscott & Ticoll, 2003): employees with less supervisor support improved in both productive and nonproductive time "because mechanics could support each other directly," while those with more supervisor support did not gain much from the new system. That speaks both to the old and new systems. Traditional performance evaluations may have been especially ineffective for employees with low supervisor support (who described them to our participant-observer as "check-the-box exercises" and "a total waste of time"), meaning they had the most to benefit from the new system. But in the new system, those who benefitted most viewed the transparent data as a way to avoid gaming behavior that can significantly, if indirectly, influence performance ratings when supervisors are responsible for performance feedback (Wayne & Liden, 1995). Even in a context like GasCo where ratings are neither subjective nor determined by a supervisor, supervisors can bias performance ratings by providing resources, plum opportunities, or recognition, while employees can bias performance ratings through a propensity to strategically ask for feedback (Wayne & Liden, 1995; DeStobbeleir et al. 2011). Transparent performance data substituted these activities for a system that one mechanic said was "just about the work."

Our relational moderator for horizontal relations, social comparison orientation, also showed a significant inverse relationship with productivity in the transition from traditional performance feedback to performance transparency. Those with a lower social comparison orientation experienced a greater productivity boost, as those who actively seek social comparison were more likely to have already sought relative performance information before the intervention and thus had less to learn from the new system. While we had left open this possibility in our hypotheses, the result is somewhat counterintuitive, as researchers tend to believe that it is the employees with high social comparison orientation—who always want to know where they stand relative to others—who will benefit most from increased performance transparency. However, in our study, those employees had less to gain because they already knew where they stood. They were also more likely to joke with the embedded participant-observer about the new system as a "competition" rather than an opportunity to "learn" or "get information," suggesting that those with high social comparison orientation may be more likely to engage in unproductive gaming activities (Charness et al., 2013; Delfgaauw et al., 2014; Hannan et al., 2012).

These findings underscore that, in traditional organizational settings, the relationships between employees and their supervisors or peers are channels, both formal and informal, by which employees may seek or receive performance feedback. To a certain degree, these preexisting channels may substitute for the effect of introducing transparent performance data and vice-versa.

Third, our study contributes to the literature on transparent performance feedback by showing that transparent performance data *in itself* can powerfully affect employee performance in the absence of strong financial incentives, career concerns, explicit goals, or supervisor intervention.

Whereas other studies show the effect of increased performance transparency combined with explicit financial incentives or in the context of strong financial incentives (Boudreau, Lakhani, & Menietti, 2016), ours empirically identifies the effect of public relative performance information *in itself* in a context of weak financial incentives and relatively little fear of career consequences. While other studies show the effect of direct supervisor observation or system-based monitoring (e.g., Blader, Gartenberg, & Prat, 2016; Griffith, 1993; Pierce et al., 2015; Staats et al., 2016), we provide empirical evidence on the effect of making that performance data transparent—known to coworkers as well as to oneself. We thus highlight the social nature of some transparency-enhancing technologies (Fulk, 1993) and empirically show some of their social-psychological effects on behavior and productivity (Bernstein, 2017).

Practical Implications

Is it productive to substitute transparent performance data for traditional performance feedback? The results from our field experiment suggest cautious optimism. On average, we find that it can boost productivity—especially by reducing nonproductive behaviors (*% Nonproductive time*)—without strong financial incentives. However, reducing negative behaviors may not increase behaviors that directly contribute to current productivity. This could be due to the lack of strong financial incentives or of a need to stand out in order to advance, in which case, adding incentives may lead to an improvement in productive behaviors. It may simply be easier to eliminate "waste" than to make the extra time productive; organizations may need to make specific efforts to accomplish that next step.

Our study also underscores the importance of relational moderators. Our findings suggest that before implementing performance transparency, one must take into account the organization's history, context, and social fabric to understand if relational mechanisms may already serve similar purposes. In short, who gets the data can determine how it is used. Relational considerations may partially explain why the revolution in transparent performance data has not always produced a revolution in productivity.

Limitations and Future Research

The key strengths of our research stem from its ability to cleanly identify the causal relationship between transparent performance data and employee performance using a carefully designed field experiment in a US service organization. However, as with any field experiment of this complexity, identifying causal relationships comes at a cost. The singular US organizational setting limits the generalizability of our findings, the single-period intervention leaves open the possibility that alternative intervention designs might affect outcomes, and the intervention's 10-week duration prevents an analysis of the longer-term effects.

Future research can extend our study beyond those limitations. In the field, there is an opportunity to explore how these findings might change—and how the relevant relational moderators might shift—in another culture where the social dynamic and attitudes towards transparency might differ or over a duration of years rather than months. In the lab, there is an opportunity to try multiple treatment conditions, designed with prior performance feedback research in mind, that go beyond our "simple" single-period intervention of showing the individual performance data in its original disaggregated form to everyone in the same work group.

Our goal in this study was to evaluate the causal impact on employee productivity of substituting transparent performance data for traditional performance feedback, so we focused less on two areas that are ripe for future research. First, our intervention—like most real-world implementations of transparent performance data—involved feedback that included more information, was more real-time, and was system generated. We therefore are unable to pinpoint which of these characteristics drives the results, which is an opportunity for future studies. Second, our focus was on understanding the transition, not improving it, and so we did not investigate moderators—such as how supervisors are trained for productive feedback sessions in the new system or other ways to improve the possibility that both traditional and new systems work well together—that might make the transition itself more successful. With this study's empirical evidence that off-the-shelf relational moderators can substantially affect such a transition, future studies in which alternative moderators serve as the core target of investigation, or even the source of the intervention itself, seem worthwhile.

The Performance Implications of Seeing Where You Stand

Because the logic of performance transparency seems so compelling, it is becoming increasingly common, either in addition to or instead of traditional supervisor-to-subordinate performance feedback. Both management scholars and managers themselves will need to take into account how the success or failure of performance transparency is affected by the specific social and interpersonal environment. We have taken an initial step in that direction. Using a field experiment, we show that transparent performance data can confer comparatively more or less motivational and learning benefit on an employee depending on how supportive his or her supervisors already are and on the degree to which that employee views his or her performance in comparison to that of coworkers. Future research—and new managerial initiatives to increase performance transparency—could be more effective with increased appreciation of the human interactions already taking place and increased attention to the relational moderators that allow us to account for them.

REFERENCES

Alvero, A. M., Bucklin, B. R., & Austin, J. 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.

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

Anderson, C. 2008. The end of theory: The data deluge makes the scientific method obsolete. ***Wired*** (June 23). http://www.wired.com/2008/06/pb-theory/. Accessed January 2, 2017.

Annett, J. 1969. ***Feedback and human behaviour: The effects of knowledge of results, incentives and reinforcement on learning and performance***. Harmondsworth, Middlesex, UK: Penguin.

Ashford, S. J. 1986. Feedback-seeking in individual adaptation: A resource perspective. ***Academy of Management Journal***, 29(3): 465–487.

Ashford, S. J., Blatt, R., & Walle, D. V. 2003. Reflections on the looking glass: A review of research on feedback-seeking behavior in organizations. ***Journal of Management***, 29(6): 773–799.

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

Ashford, S. J., & Tsui, A. S. 1991. Self-regulation for managerial effectiveness: The role of active feedback seeking. ***Academy of Management Journal***, 34(2): 251–280.

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

Balzer, W. K., Doherty, M. E., & O'Connor, R. 1989. Effects of cognitive feedback on performance. ***Psychological Bulletin***, 106(3): 410–433.

Bandura, A. 1991. Social cognitive theory of self-regulation. ***Organizational Behavior and Human Decision Processes***, 50(2): 248–287.

Baron, R. M. & Kenny, D. A. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. ***Journal of Personality and Social Psychology***, 51(6): 1173-1182.

Becker, H. S. 1958. Problems of inference and proof in participant observation. ***American Sociological Review***, 23(6): 652–660.

Becker, L. J. 1978. Joint effect of feedback and goal setting on performance: A field study of residential energy conservation. ***Journal of Applied Psychology***, 63(4): 428–433.

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

Bernstein, E. 2014. The transparency trap. ***Harvard Business Review***, 92(10): 58–66.

Bernstein, E. 2017. Making transparency transparent: The evolution of observation in management theory. ***Academy of Management Annals***, 11(1): 217-266.

Bernstein, E., & Blunden, H. 2015. The sales director who turned work into a fantasy sports competition. ***Harvard Business Review*** (March 27). https://hbr.org/2015/03/the-sales-director-who-turned-work-into-a-fantasy-sports-competition. Accessed January 2, 2017.

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

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Milkman, K. L. 2015. The effect of providing peer information on retirement savings decisions. ***Journal of Finance***, 70(3): 1161–1201.

Blader, S., Gartenberg, C., & Prat, A. 2016. ***The contingent effect of management practices.*** Columbia Business School Research Paper No. 15-48, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2594258. Accessed January 2, 2017.

Boudreau, K. J., Lakhani, K. R., & Menietti, M. 2016. Performance responses to competition across skill levels in rank-order tournaments: Field evidence and implications for tournament design. ***RAND Journal of Economics***, 47(1): 140–165.
Bradler, C., Dur, R., Neckermann, S., & Non, A. 2013. Employee recognition and performance: A field experiment. ***ZEW-Centre for European Economic Research Discussion Paper 13-017***.
Brown, J. L., Fisher, J. G., Sooy, M., & Sprinkle, G. B. 2014. The effect of rankings on honesty in budget reporting. ***Accounting, Organizations and Society***, 39(4): 237–246.
Buckingham, M., & Goodall, A. 2015. Reinventing performance management. ***Harvard Business Review***, 93(4): 40–50.
Charness, G., Masclet, D., & Villeval, M. C. 2013. The dark side of competition for status. ***Management Science***, 60(1): 38–55.
Cialdini, R. B., & Trost, M. R. 1998. Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), ***The Handbook of Social Psychology***, 4th Edition, Chapter 21, p. 151-192. New York, NY: McGraw-Hill.
Culbert, S., & Rout, L. 2010. ***Get rid of the performance review! How companies can stop intimidating, start managing—and focus on what really matters.*** New York: Hachette.
Cunningham, L., & McGregor, J. 2015. Why big business is falling out of love with the annual performance review. ***Washington Post*** (August 17). https://www.washingtonpost.com/news/on-leadership/wp/2015/08/17/why-big-business-is-falling-out-of-love-with-annual-performance-reviews/. Accessed January 2, 2017.
Delfgaauw, J., Dur, R., Non, A., & Verbeke, W. 2014. Dynamic incentive effects of relative performance pay: A field experiment. ***Labour Economics***, 28: 1–13.
DeNisi, A. S., & Kluger, A. N. 2000. Feedback effectiveness: Can 360-degree appraisals be improved? ***Academy of Management Executive***, 14(1): 129–139.
DeStobbeleir, K. E., Ashford, S. J., & D. Buyens. 2011. Self-regulation of creativity at work: The role of feedback-seeking behavior in creative performance. ***Academy of Management Journal***, 54(4) 811–831.
Dockstader, S. L., Nebeker, D. M., & Shumate, E. C. 1977. ***The effects of feedback and an implied standard on work performance***. San Diego, CA: ERIC.
Dunn, J., Ruedy, N. E., & Schweitzer, M. E. 2012. It hurts both ways: How social comparisons harm affective and cognitive trust. ***Organizational Behavior and Human Decision Processes***, 117(1): 2–14.
Eisenberger, R., Stinglhamber, F., Vandenberghe, C., Sucharski, I. L., & Rhoades, L. 2002. Perceived supervisor support: Contributions to perceived organizational support and employee retention. ***Journal of Applied Psychology***, 87(3): 565–573.
Emerson, R. M., Fretz, R. I., & Shaw, L. L. 1995. ***Writing ethnographic fieldnotes***. Chicago: University of Chicago Press.
Endler, N. S. 1993. Personality: An interactional perspective. In J. Hettema & I. J. Deary (Eds.), ***Foundations of Personality***: 251–268. Dordrecht, Netherlands: Kluwer.
Fedor, D. B. 1991. Recipient responses to performance feedback: A proposed model and its implications. ***Research in Personnel and Human Resources Management***, 9: 73–120.
Festinger, L. 1954. A theory of social comparison processes. ***Human Relations***, 7(2): 117–140.
Florin-Thuma, B. C., & Boudreau, J. W. 1987. Performance feedback utility in a small organization: Effects on organizational outcomes and managerial decision processes. ***Personnel Psychology***, 40(4): 693–713.
Fulk, J. 1993. Social construction of communication technology. ***Academy of Management Journal***, 36(5): 921–950.

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

Goldstein, J. 2014. To increase productivity, UPS monitors drivers' every move. ***NPR Planet Money Morning Edition*** (April 17).

Grant, A. M., & Parker, S. K. 2009. Redesigning work design theories: The rise of relational and proactive perspectives. ***Academy of Management Annals***, 3(1): 317–375.

Griffith, T. L. 1993. Monitoring and performance: A comparison of computer and supervisor monitoring. ***Journal of Applied Social Psychology***, 23(7): 549–572.

Guzzo, R. A., Jette, R. D., & Katzell, R. A. 1985. The effects of psychologically based intervention programs on worker productivity: A meta-analysis. ***Personnel Psychology***, 38(2): 275–291.

Haas, J. R., & Hayes, S. C. 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., McPhee, G. P., Newman, A. H., & Tafkov, I. D. 2012. The effect of relative performance information on performance and effort allocation in a multi-task environment. ***The Accounting Review***, 88(2): 553–575.

Head, S. 2014. ***Mindless: Why smarter machines are making dumber humans***. New York: Basic.

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

Huhman, H. R. 2014. 5 ways HR technology can improve performance reviews. ***Entrepreneur*** (July 23). https://www.entrepreneur.com/article/235797. Accessed January 2, 2017.

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

Ilgen, D. R., Peterson, R. B., Martin, B. A., & Boeschen, D. A. 1981. Supervisor and subordinate reactions to performance appraisal sessions. ***Organizational Behavior and Human Performance***, 28(3): 311–330.

Ivancevich, J. M., & McMahon, J. T. 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.

Jaccard, J. & Turrisi, R. 2003. ***Interaction effects in multiple regression (2nd ed): Quantitative Applications in the Social Sciences***. Thousand Oaks, CA: Sage.

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

Jorgensen, D. L. 1989. ***Participant observation: A methodology for human studies***. Thousand Oaks, CA: Sage.

Kantor, J., & Streitfeld, D. 2015. Inside Amazon: Wrestling big ideas in a bruising workplace. ***New York Times*** (August 15). http://www.nytimes.com/2015/08/16/technology/inside-amazon-wrestling-big-ideas-in-a-bruising-workplace.html?_r=0. Accessed January 2, 2017.

Kc, D. S., & Terwiesch, C. 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.

Kidwell, R. E., & Sprague, R. 2009. Electronic surveillance in the global workplace: Laws, ethics, research and practice. ***New Technology, Work and Employment***, 24(2): 194–208.

Kluger, A. N., & DeNisi, A. 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., Lewinsohn, S., & Aiello, J. R. 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.

Komaki, J., Barwick, K. D., & Scott, L. R. 1978. A behavioral approach to occupational safety: Pinpointing and reinforcing safe performance in a food manufacturing plant. ***Journal of Applied Psychology***, 63(4): 434–445.

Kottke, J. L., & Sharafinski, C. E. 1988. Measuring perceived supervisory and organizational support. ***Educational and Psychological Measurement***, 48(4): 1075–1079.

Kruger, J., & Dunning, D. 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.

Larkin, I., Pierce, L., & Gino, F. 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.

Latham, G. P., & Locke, E. A. 1991. Self-regulation through goal setting. ***Organizational Behavior and Human Decision Processes***, 50(2): 212–247.

Lewis, H. B. 1971. Shame and guilt in neurosis. ***Psychoanalytic Review***, 58(3): 419–438.

Locke, E. A., & Latham, G. P. 1990. ***A theory of goal setting & task performance.*** Englewood Cliffs, NJ: Prentice Hall.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. 1981. Goal setting and task performance: 1969-1980. ***Psychological Bulletin***, 90(1): 125–152.

Lount, Jr., R. B., & Wilk, S. L. 2014. Working harder or hardly working? Posting performance eliminates social loafing and promotes social laboring in workgroups. ***Management Science,*** 60(5): 1098-1106.

Lublin, J. S. 2011. Transparency pays off in 360-degree reviews. ***Wall Street Journal*** (December 8). http://www.wsj.com/articles/SB10001424052970203501304577086592075136080. Accessed January 2, 2017.

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

McFarland, C., & Miller, D. T. 1994. The framing of relative performance feedback: Seeing the glass as half empty or half full. ***Journal of Personality and Social Psychology***, 66(6): 1061–1073.

Michinov, E., & Michinov, N. 2001. The similarity hypothesis: A test of the moderating role of social comparison orientation. ***European Journal of Social Psychology***, 31: 549-555.

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

Northcraft, G. B., & Ashford, S. J. 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.

Pfeffer, J. 2016. Sorry, Uber. Customer service ratings cannot replace managers. ***Fortune*** (March 31). http://fortune.com/2016/03/31/sorry-uber-customer-service-ratings-cannot-replace-managers/. Accessed January 2, 2017.

Pierce, L., Snow, D., & McAfee, A. 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., & Farh, J. L. 1989. Effects of feedback sign and credibility on goal setting and task performance. ***Organizational Behavior and Human Decision Processes***, 44(1): 45–67.

Pritchard, R. D., Bigby, D. G., Beiting, M., Coverdale, S., & Morgan, C. 1981. ***Enhancing productivity through feedback and goal setting***. DTIC Document.

Rawlinson, K. 2013. Tesco accused of using electronic armbands to monitor its staff. ***The Independent*** (February 13). http://www.independent.co.uk/news/business/news/tesco-accused-of-using-electronic-armbands-to-monitor-its-staff-8493952.html. Accessed January 2, 2017.

Ren, Y., & Argote, L. 2011. Transactive memory systems 1985-2010: An integrative framework of dimensions, antecedents, and consequences. ***Academy of Management Annals***, 5(1): 189–229.

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

Roethlisberger, F. J., & Dickson, W. J. 1939. ***Management and the worker: An account of a research program conducted by the Western Electric Company***. Cambridge, MA: Harvard University Press.

Sansone, C. 1986. A question of competence: The effects of competence and task feedback on intrinsic interest. ***Journal of Personality and Social Psychology***, 51(5): 918–931.

Schultz, K. L., Juran, D. C., & Boudreau, J. W. 1999. The effects of low inventory on the development of productivity norms. ***Management Science***, 45(12): 1664–1678.

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

Song, H., Tucker, A., Murrell, K. L., & Vinson, D. R. 2016. ***Public relative performance feedback in complex service systems: Improving productivity through the adoption of best practices***. Working paper 16-043, Harvard Business School, Boston, MA.

Spradley, J. P. 1980. ***Participant observation***. Belmont, CA: Wadsworth.

Staats, B. R., Dai, H., Hofmann, D., & Milkman, K. L. 2016. Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. ***Management Science***.

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

Suls, J., Martin, R., & Wheeler, L. 2002. Social comparison: Why, with whom, and with what effect? ***Current Directions in Psychological Science***, 11(5): 159–163.

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

Taylor, G. 1985. ***Pride, shame, and guilt: Emotions of self-assessment***. Gloucester, UK: Clarendon.

Taylor, M. S., Fisher, C. D., & Ilgen, D. R. 1984. Individual's reactions to performance feedback in organizations: A control theory perspective. In K. M. Rowland & G. R. Ferris (Eds.), ***Research in personnel and human resources management***, vol. 2: 81–124. Greenwich, CT: JAI.

Thorndike, E. L. 1927. The law of effect. ***American Journal of Psychology***, 39(1/4): 212–222.

Van Quaquebeke, N., van Knippenberg, D., & Eckloff, T. 2011. Individual differences in the leader categorization to openness to influence relationship: The role of followers' self-perception and social comparison orientation. ***Group Processes & Intergroup Relations***, 14(5): 605-622.

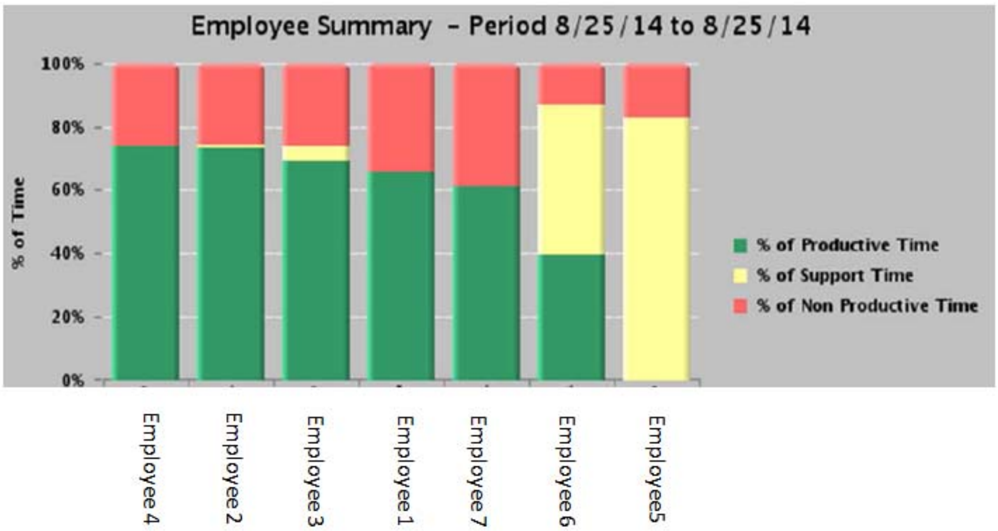
Vroom, V. H. 1964. ***Work and motivation***. New York: Wiley.

Wayne, S. J., & Liden, R. C. 1995. Effects of impression management on performance ratings: A longitudinal study. ***Academy of Management Journal***, 38(1) 232–260.

Wegner, D. M. 1987. Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), ***Theories of group behavior***: 185–208. New York: Springer-Verlag.

Wood, J. V. 1989. Theory and research concerning social comparisons of personal attributes. ***Psychological Bulletin***, 106(2): 231–248.

Figure 1
Screenshot of the Mechanic's View of the Daily Scorecard Information



Productive Time

- En route
- Onsite

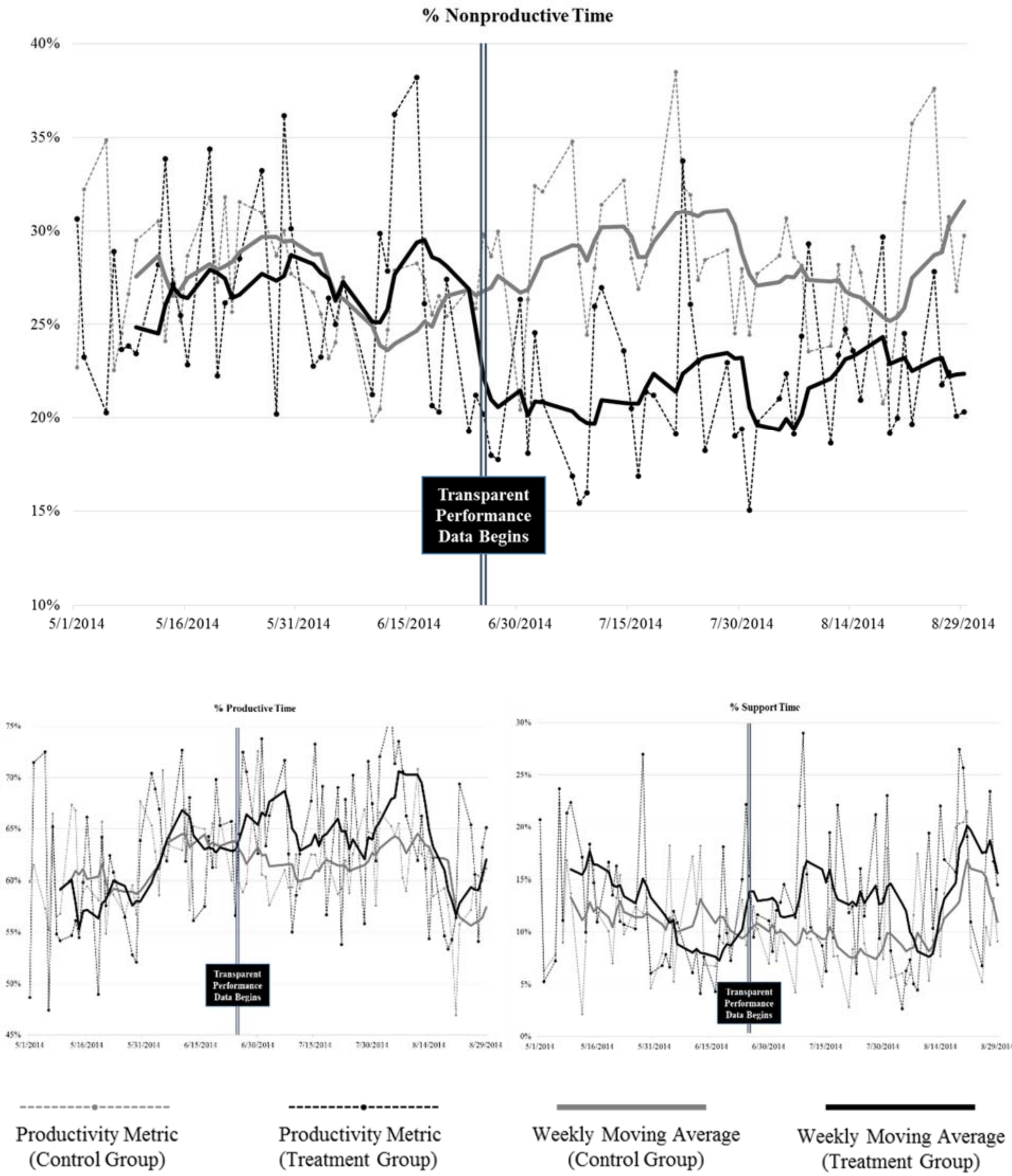
Support Time

- 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 2
Visualization of the Data



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 respectively). 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	Obs.	Mean	SD	Median
% Nonproductive Time	A mechanic's nonproductive hours as a percentage of the total available working hours.	11,120	30.70	23.41	23.47
% Productive Time	A mechanic's productive hours as a percentage of the total available working hours.	11,120	59.13	24.36	66.11
Tenure	Number of years the mechanic had worked in the company at the start of the intervention.	11,120	18.52	8.28	19.00
Age	The mechanic's age (in years) at the beginning of the intervention.	11,120	47.21	8.18	50.00
White	=1 if the ethnic group of the mechanic is White.	11,120	0.55	0.50	1.00
Social Comparison Orientation	Sum of the scores for the 11 questions on social comparison orientation in survey.	11,120	35.19	7.65	37.00
Prior Performance Nonproductive	The average "% Nonproductive Time" in the pre-intervention period.	11,120	30.64	12.56	28.64
Prior Performance Productive	The average "% Productive Time" in the pre-intervention period.	11,120	59.55	12.39	62.94
Self-evaluation	The self-reported percentile of the pre-intervention performance on % Productive Time.	11,120	67.33	31.71	81.89
Supervisor Support	Sum of the scores for the questions on supervisor support in the pre-experimental survey.	11,120	24.16	4.01	24.00

Table 2
Correlations

Variable	1	2	3	4	5	6	7	8	9
1. % Nonproductive Time									
2. % Productive Time	**-0.77**								
3. Tenure	-0.00	**-0.04**							
4. Age	**0.04**	**-0.03**	**0.71**						
5. White	**0.07**	**-0.14**	**0.03**	**-0.30**					
6. Social Comparison Orientation	**0.03**	**0.03**	**-0.18**	**-0.05**	**-0.06**				
7. Prior Performance Nonproductive	**0.54**	**-0.45**	0.01	**0.10**	**0.13**	**0.07**			
8. Prior Performance Productive	**-0.48**	**0.51**	**-0.08**	**-0.07**	**-0.26**	**0.04**	**-0.89**		
9. Self-evaluation	**-0.12**	**0.05**	**-0.19**	**-0.09**	**-0.16**	0.02	**-0.20**	**0.12**	
10. Supervisor Support	**-0.05**	**0.04**	**-0.19**	**-0.22**	**0.11**	**0.48**	**-0.10**	**0.09**	**0.10**

n = 11,120; all correlations above or equal to 0.03 are significant at p < .01 (as bolded).

Table 3
Does Performance Transparency Improve Productivity?

	(1) % Non-productive Time	(2) % Productive Time	(3) % Non-productive Time	(4) % Productive Time	(5) % Non-productive Time	(6) % Productive Time
Treat x Post	-1.87*	0.13	-3.38***	2.82	-3.40**	2.70
	(1.07)	(2.29)	(1.09)	(3.84)	(1.10)	(3.35)
Treat	-0.11	0.41	0.11	0.07		
	(0.25)	(0.39)	(0.33)	(0.33)		
Post	0.57	-0.88	1.23	-3.50***	1.35	-3.59***
	(0.99)	(1.04)	(0.99)	(0.83)	(1.06)	(0.72)
Tenure	0.04	-0.02	-0.05	-0.03		
	(0.04)	(0.04)	(0.03)	(0.03)		
Age	-0.07	-0.03	-0.06	0.022		
	(0.05)	(0.06)	(0.11)	(0.03)		
White	-0.13	-0.98	-0.33	-0.12		
	(0.35)	(0.97)	(0.42)	(0.38)		
Social Comparison			-0.05	0.06		
			(0.05)	(0.05)		
Prior Performance		0.92***		0.97***		
(Productive)		(0.06)		(0.01)		
Prior Performance	0.97***		0.99***			
(Nonproductive)	(0.03)		(0.01)			
Self-evaluation			-0.01	-0.00		
			(0.01)	(0.01)		
Supervisor Support			0.06	-0.13		
			(0.11)	(0.13)		
Indiv. fixed effects?	No	No	No	No	Yes	Yes
Observations	26,993	26,993	11,120	11,120	11,120	11,120
Adj. Rsq	0.26	0.02	0.29	0.26	0.15	0.16

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of % Nonproductive Time (Columns 1, 3, and 5) and % Productive Time (Columns 2, 4, and 6) on a treatment indicator ("Treat"), a post-intervention indicator ("Post"), an interaction of the two variables ("Treat x Post"), and other controls.

Columns 1 and 2 are analyses conducted on a larger sample of mechanics for whom we only have demographic and performance data. Columns 3 and 4 are analyses conducted on the sample for whom we also have survey data (and hence can construct measures for social comparison orientation, perceived level of supervisor support, and self-evaluation). For consistency, we will use the latter sample for all subsequent analyses. Columns 5 and 6 serve as a robustness check by incorporating individual fixed effects for the latter sample. In Columns 5 and 6, the "Treat" indicator is absorbed by individual fixed effects and is not reported. All variables are defined in Table 1. Standard errors are block-bootstrapped, clustered at the work-center level, and reported in parentheses.

Table 4
Supervisor Support as a Moderator for the Impact of Substituting Performance Transparency for Traditional Performance Feedback

	(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	-2.20	-5.74***	0.90	9.72***
	(1.43)	(1.89)	(4.85)	(1.19)
Treat	0.25	0.40	-0.99	-0.31
	(0.35)	(2.08)	(0.48)	(2.21)
Post	0.16	3.03***	-3.47**	-3.63***
	(1.35)	(1.15)	(1.39)	(0.79)
Tenure	0.00	-0.06	-0.03	-0.07
	(0.05)	(0.07)	(0.05)	(0.09)
Age	-0.04	-0.09	0.02	0.06
	(0.05)	(0.09)	(0.06)	(0.10)
White	-0.41	0.18	0.22	-0.59
	(0.66)	(0.80)	(0.51)	(0.80)
Social Comparison	0.01	-0.05	0.05	0.050
	(0.06)	(0.10)	(0.05)	(0.12)
Prior Performance			1.01***	0.97***
(Productive)			(0.03)	(0.04)
Prior Performance	1.01***	0.98***		
(Nonproductive)	(0.03)	(0.04)		
Self-evaluation	0.00	-0.02	-0.01**	0.01
	(0.01)	(0.02)	(0.01)	(0.02)
Supervisor Support	-0.03	0.30	-0.05	-0.27
	(0.09)	(0.38)	(0.08)	(0.42)
Observations	7,159	3,961	7,159	3,961
Adj. Rsq	0.23	0.36	0.24	0.30

*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 treatment indicator (“Treat”), a post-intervention indicator (“Post”), an interaction of the two variables (“Treat x Post”), and other controls. “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 block-bootstrapped, clustered at the work-center level, and reported in parentheses.

Table 5
Social Comparison Orientation as a Moderator for the Impact of Substituting Performance Transparency for Traditional Performance Feedback

	(1) % Nonproductive Time High Social Comparison	(2) % Nonproductive Time Low Social Comparison	(3) % Productive Time High Social Comparison	(4) % Productive Time Low Social Comparison
Treat x Post	-2.73	-5.33***	5.04	2.43
	(1.96)	(1.50)	(3.38)	(5.18)
Treat	0.67	-0.06	-0.29	0.28
	(0.67)	(0.63)	(1.28)	(1.02)
Post	-0.28	3.74**	-2.43**	-5.22***
	(1.21)	(1.52)	(1.02)	(1.65)
Tenure	-05.56	0.06	0.02	-0.10
	(0.05)	(0.05)	(0.06)	(0.08)
Age	-0.03	-0.03	-0.01	0.055
	(0.06)	(0.06)	(0.06)	(0.10)
White	-0.64	-0.02	0.44	-0.47
	(0.67)	(1.04)	(0.80)	(1.19)
Social Comparison	-0.88	0.05	0.08	-0.01
	(0.07)	(0.06)	(0.14)	(0.06)
Prior Performance			0.98***	0.99***
(Productive)			(0.03)	(0.06)
Prior Performance	1.00***	0.99***		
(Nonproductive)	(0.03)	(0.02)		
Self-evaluation	-0.01	-0.02	-0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Supervisor Support	0.10	-0.08	-0.14	-0.15
	(0.10)	(0.08)	(0.12)	(0.15)
Observations	5,909	5,211	5,909	5,211
Adj. Rsq	0.29	0.29	0.26	0.25

*p < .10, **p < .05, ***p < .01. This table reports OLS regression results of % Nonproductive Time (Columns 1 and 2) and % Productive Time (Columns 3 and 4) on a treatment indicator (“Treat”), a post-intervention indicator (“Post”), an interaction of the two variables (“Treat x Post”), and other controls. “High Social Comparison” (“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 block-bootstrapped, clustered at the work-center level, and reported in parentheses.