Promoting Core Values with a Values-based Performance Assessment System: Evidence from a Retail Chain

Carolyn Deller
The Wharton School
University of Pennsylvania
cdeller@wharton.upenn.edu

Susanna Gallani (*)
Harvard Business School
sgallani@hbs.edu

Tatiana Sandino
Harvard Business School
tsandino@hbs.edu

August 2024

Abstract:

We analyze the outcomes of introducing a formal control system—a values-based performance assessment system—to disseminate and reinforce corporate values. Using data from a retail chain, we examine the system's impact on variables capturing behaviors consistent with those values. The system's effectiveness exhibited significant variation. Specifically, we find it to be more effective in stores with greater promotion opportunities—consistent with the system motivating increased effort where implicit incentives are stronger—and in stores with managers who had experience related to the implementation of the promoted values—consistent with the system cultivating skills necessary for acting on core values. However, the system was counterproductive along various dimensions where management's actions were likely to be perceived as inconsistent with the espoused values. Our findings highlight important factors for instilling core values via formal management control systems.

Keywords: 360-degree assessments, values, implicit incentives, performance evaluation.
JEL codes: M12, M40
Data availability: The data used in this project was provided to the authors on a proprietary basis and cannot be shared without the express consent of the organization's legal representatives.
Responsible Conduct of Research: This project was reviewed and approved by the Harvard University Area Internal Review Board (HUA IRB). Approval documentation is available upon request by contacting Alain Bonacossa (abonacossa@hbs.edu) and referencing case #IRB15-0752.

(*) Corresponding author: Susanna Gallani, Harvard Business School, Soldiers Field Boston, MA

For useful comments and suggestions, we thank Markus Arnold, Pablo Casas-Arce, Tony Davila, John Harry Evans, Mario Schabus, and participants at the 2019 Information, Markets, and Organizations Conference at Harvard Business School, the 2019 AIMA World Conference on Management Accounting Research, the 2019 Consortium for Operational Excellence in Retailing, the 2019 Cherry Blossom Conference at George Washington University, the 2018 AAA Annual Meeting, the 2018 Emerging Management Accounting Scholars Symposium, the 2018 Global Management Accounting Research Symposium, the 2018 Management Accounting Section Midyear Meeting, University of Central Florida, Erasmus University, and Harvard Business School. We thank Kyle Thomas for his excellent assistance at the research site and in the data collection and analyses, Mayuresh Kumar and Gunjan Kapil for their research assistance at the research site, Albert Shin and Trang Nguyen for research assistance later in the project, and Andrew Marder and Xiang Ao for helpful advice on econometric issues. We also thank the mobile phone retail company where we conducted our study for collaborating with us on this project and providing us with access to data. All errors are our own.

1. INTRODUCTION

Employee alignment with corporate values—essential to attain corporate goals—can be fostered by (a) hiring people whose values naturally align with the organization's core values, (b) shaping employees' values, and/or (c) firing employees who violate those values (Chatman, 1991). We define corporate values—alternatively referred to as "core values," "organizational values," or simply "values"—as the set of principles that underlie an organization's operational foundation and direct its employees' attitudes and behaviors at work (Lencioni, 2002; Silva, 2021). An emerging literature in management accounting on the role of formal management control systems in disseminating corporate values has focused primarily on the role of recruiting employees naturally aligned with organizational values, paying less attention to formal management control systems used to influence employees' values and monitor their adherence to them. Anecdotal evidence suggests that formal management control systems can drive value alignment (see for example, the Henkel teaching case, in which Simons and Kindred (2012) document values-based performance assessments driving a company's redefinition of its core values). While evidence suggests that organizations can use informal methods to instill values (e.g., Chatman, 1991; Martin, 2016), the degree to which formal control systems, such as performance assessment systems, can do so is less clear.

We examine a *values-based performance assessment system*; that is, one designed to formalize, communicate, monitor, and reinforce an organization's core values. This is a commonly used formal management control system. For instance, in a WorldatWork survey of 254 human resource managers from a diverse set of firms, 72% indicated that their company's performance evaluation process reflected organizational core values to a significant extent.[1] We study whether and under what conditions the implementation of a values-based performance assessment system (hereafter, values-based PA system) leads to changes in performance and behaviors consistent with the promoted values.

[1] This is the percentage of respondents indicating that the extent to which their company's performance evaluation process reflects guiding principles/corporate values was 5 or greater on a Likert scale ranging from 1 ("not at all") to 7 ("to a large extent"). WorldatWork conducted this survey in collaboration with several researchers cited in the reference list (WorldatWork, 2019). The survey includes organizations from multiple industries and of various sizes.

The management literature identifies two theoretical mechanisms by which performance assessments could effectively promote desired behaviors (e.g., Boswell & Boudreau, 2002). First, the values-based PA system could operate through an *evaluative* mechanism providing implicit incentives, especially to employees who find the values inspiring and are interested in pursuing a career in the organization. This mechanism can drive greater effort in activities and behaviors that embody corporate values. Second, the system could foster value alignment through a *developmental* mechanism through which employees (both raters and ratees) learn—or learn more—about desired behaviors that create value and that make their work more meaningful to themselves and more valuable to the organization. This understanding should enhance employees' skills and empower them to prioritize activities that uphold corporate values.

On the flip side, theory predicts that a values-based PA system might not align employee behaviors with corporate values due to the tacit knowledge and subjective criteria required to discern desirable and undesirable behaviors (Castillo, 2002). The difficulty of translating values into a formal measurement system could undermine such a system. Furthermore, the system could have unintended negative consequences if it raises employees' expectations of their leaders' adherence to corporate values. Cha and Edmondson (2006) theorize that employees may become disillusioned if they perceive that their leaders either failed to behave according to core values or failed to support employees' efforts to uphold those values. We examine the overall effectiveness of a values-based PA system in shaping employee behavior and, more importantly, the factors that may enhance or undermine the underlying mechanisms.

We analyze data from 32 stores in a growing retail chain in a major city in India. When the company was relatively small, the managing director personally shaped and monitored workers' behavior through direct informal communication of her vision and values by working alongside store teams and mentoring them as they interacted with customers. However, as the company scaled up, management decided to develop a formal system (specifically, a values-based PA system) to disseminate and reinforce the company's vision and core values.[2]

[2] While designing and introducing this system, the managing director solicited advice from our research team. She led the pilot implementation in randomly selected stores and agreed to share company data with us to analyze its effects.

The company implemented a values-based 360-degree system to formalize and communicate to all workers the four values that mattered most to the company: (1) working hard to gain control of one's own career; (2) giving more value; (3) being honest and ethical, and (4) being caring and respectful. The system was designed to evaluate the alignment between the store manager's behavior and the organization's core values and was based on surveys completed by the store managers, their team members, and their supervisors. The questions were grouped by core value, capturing behaviors desired by senior management. Store managers were expected to help their teams adopt behaviors aligned with corporate values. However, by enabling the raters (i.e., the store manager's supervisors and store staff members) to evaluate the congruence of the store manager's behaviors with those values, the system also communicated corporate values and highlighted desired behaviors to the store staff members. A secondary goal was to communicate what behaviors mattered for career advancement and to identify workers suitable for promotions. A separate existing monetary incentive plan linked to financial goals was kept and the new values-based system did not come with any new monetary incentives.

To test the effectiveness of the values-based PA system, the company implemented it in a randomly selected treatment group comprising half the chain's stores, in consultation with our research team. We conduct a difference-in-differences analysis of the outcomes of this intervention using 918 store-weeks from 32 stores, though the sample is reduced in some of our tests by the availability of some variables. We estimate the effect of the system using proxies of the workers' alignment with the company's four core values. More importantly, we examine four factors informed by theory potentially affecting the system's effectiveness. The first two affect the evaluative and developmental mechanisms expected to drive performance assessment effectiveness: (1) heterogeneity in *promotion opportunities* across units (in this case, stores), potentially affecting the strength of the evaluative mechanism associated with the system and consequently the workers' motivation to exert greater effort to observe company values; and (2) variation in the *unit managers' experience* (in this case, regarding how to behave according to corporate values), potentially affecting the strength of the system's developmental mechanism and consequently the workers' ability to adhere to all core values. The next two conditions increase the probability that some workers

would become disillusioned with leadership, perceiving the system to be duplicitous and rendering it harmful instead of effective; namely, (3) variation in the risk that the leaders' behavior is perceived to lack congruency with the core values promoted and (4) variation in the risk that the leaders' support for the core values is perceived to be inadequate (e.g., in teams receiving insufficient inventory from leaders to *give more value* to customers). Examining such heterogeneity is important, as average effects can mask heterogeneous effects that are important for managers to consider (Bryan et al., 2021).

Our sample period ends *before* any survey results or feedback were provided to store managers at treated stores. Thus, rather than documenting the ex-post feedback effects of a values-based PA system, we examine its effects on the behavior of workers who are (a) formally learning about the core values, (b) assessing desired behaviors aligned with those values, and (c) anticipating being evaluated on those values in the future.[3] Our design allows us to study whether the values-based PA system has an effect on the workers' behaviors by formalizing the organization's core values and making them salient to the workers.

Our main dependent variables are core-values–related store-level metrics capturing behaviors and outcomes that could be influenced by the store managers and their staff within a short timeframe—sales, customer service, honesty, and caring/respectful treatment of others. We find no significant average effects attributable to the intervention on any of the available measures associated with the company's core values. However, we identify conditions affecting the values-based PA system's effectiveness.

First, the intervention had a more positive effect in stores where the *evaluative* aspect of the system was likely to be more salient. Specifically, stores with greater opportunities for promotion responded to the intervention with (a) larger increases in sales, (b) a higher likelihood of customers recommending the retail chain, and (c) a more frequent provision of complete solutions for customers. The increase in sales is consistent with the values-based PA system promoting "working hard" as a core value, while the other two results align with the system promoting "giving more value" to customers as a core value.

[3] Since the values-based PA system assessed the store managers' behaviors, store managers expected to be evaluated again in the future based on the system, while other raters might have expected that these values would be important criteria for their career progression.

Second, the values-based PA system had a more positive effect on behaviors associated with the core values related to "working hard" and "being caring and respectful" where its *developmental* aspect was likely to be more effective. That is, we find that the effects on sales as well as on customer ratings of politeness and helpfulness were amplified in stores with managers that had sufficient experience with core values (and thus greater ability to prioritize, model, and facilitate value-related behaviors).

Our study also reveals conditions leading to a less effective or even harmful intervention in certain stores, which may help explain the intervention's insignificant average effect. Specifically, store teams seemed less receptive to the values-based PA system where there was a higher risk of disillusionment with the values exhibited by company leadership or the support provided by them. Some store staff members questioned the leaders' adherence to the company's core values of "being honest and ethical" given that they weren't always paid on time. In addition, while the core values emphasized "giving more value" to customers, some store staff claimed that management did not provide them with enough inventory to do so. In line with theory, these concerns suggest the PA system raised workers' scrutiny of their leaders' behaviors and their congruence with core values. We leverage company data to analyze whether the intervention's effectiveness depended on the timing of compensation payments and inventory levels. We find that treated stores whose managers had been paid late before the intervention responded more negatively, as evidenced by a lower likelihood of customers recommending the retail chain, following the implementation of the values-based PA system. Regarding inventory levels, treated stores in the lowest quintile of pre-intervention average days-of-sales in inventory had a marginally lower percentage of invoices including bundles, that is, they were less likely to provide more complete solutions for customers, following the intervention. These results suggest potentially disillusioned employees exerted less effort on "giving more value" to customers. We conclude that a values-based PA system is likely to demotivate rather than motivate workers to incorporate core values into everyday operations if they perceive management violates those values. However, by helping to unearth leadership behaviors that are perceived to be inconsistent with core values in the short run, a values-based PA system could lead to improvements in management practices in the long run. For example, insights from this study led management at our research

site to address problems with late payments and understocked stores, ultimately fostering a better environment in which to promulgate its core values.

We make several contributions to the literature and practice of management accounting. First, we contribute to an emerging literature on the role of formal management controls in disseminating corporate values. Campbell and Sandino (2019) conceptually discuss the role of formal control systems in strengthening core values, including systems used to (a) recruit employees fitting those values (e.g., formal fit-focused selection systems), (b) orient employee beliefs towards core values (e.g., training and assessment systems promoting core values), and (c) prevent misalignment over time (i.e., reinforcement systems such as symbols and traditions highlighting core values). Only a few management control studies have empirically tested the effect of formal control systems on the dissemination and enforcement of core values and most of these focus on recruiting systems. For instance, Liu, Liu, and Chu (2019) and Cai (2023) find that formal, fit-focused selection systems can improve employee retention (attributed to better value alignment) and performance. Our results show that, under certain circumstances, a formal values-based PA system can help workers learn about core organizational values and put them into action. We also highlight the relevance of management by example, suggesting that a system emphasizing core values is unlikely to be effective if its leaders are not perceived to be exemplifying those values themselves.

Second, we extend academic knowledge on the effective use of performance assessment systems. To our knowledge, our study is the first to directly examine whether and when performance assessment systems can be used to promulgate behaviors consistent with core values. Values-based performance assessment systems differ from other assessment systems in that they (a) aim to shape behaviors that are hard to measure comprehensively, given that they are largely tacit and subjective, (b) seek to propagate values that have the potential to make work more meaningful for employees, and (c) can change employee expectations with respect to the behaviors that should be adopted not only by themselves but by their leaders. Rather than examining the effect of implementing a values-based PA system on overall performance ratings (as do other performance assessment studies), we examine its effect on variables capturing changes in the workers' adherence to core values. This allows us to assess behavioral changes from before to after the system was

implemented (whereas performance ratings from the system itself would not enable such comparisons). We also examine conditions affecting the effectiveness of a values-based PA system. Our results suggest that such a system is likely to be more effective when workers (including raters and ratees) have (a) higher implicit incentives due to promotion opportunities and (b) greater ability to implement values due to their or their managers' experience with core values. Conversely, a values-based PA system is likely to be less effective when workers may perceive that top management violates the espoused values or does not support them in pursuing those values.

Third, our work contributes to a broader management control literature on the effects of control mechanisms unrelated to monetary incentives, such as transparent performance feedback (Bernstein and Li 2017), relative performance information (Blanes i Vidal & Nossol, 2011; Casas-Arce, Deller, Martinez-Jerez, & Narciso, 2021), and recognition (Bradler, Dur, Neckermann, & Non, 2016; Gallani, 2022). A values-based PA system aims to drive employee motivation and organizational alignment by shaping a more consistent and purposeful organizational culture.

2. LITERATURE REVIEW, HYPOTHESIS, AND RESEARCH QUESTIONS

The accounting and management literatures have long suggested that organizations rely on culture to foster employee alignment with organizational goals and values. Formalizing values not only facilitates such alignment but may also help employees find more meaning in their roles and thus enhance their motivation (Rosso, Dekas, & Wrzesniewski, 2010; Kim, Chang, & Kim, 2018). These literatures recommend recruiting and socializing employees in a manner that would lead them to identify with the organization's values and goals (Ouchi, 1979; Chatman, 1991; Rosso, Dekas, & Wrzesniewski, 2010). Research has suggested that informal management control practices—including social interactions, mentorship, and selection through employee referrals—could help instill organizational values (Chatman, 1991; Campbell, 2012). More recently, the accounting and management literatures have recognized that formal control systems could also shape values across the organization (Campbell & Sandino, 2019).

An emerging literature in management accounting has found the implementation of formal, fit-focused selection systems (Liu et al., 2019; Cai, 2023) is associated with longer employee tenure and improved performance, both attributed to better alignment with company values. However, this literature has paid less attention to the use of formal management control systems to instill and reinforce corporate values among *incumbent* employees. In practice, numerous organizations—particularly those that are growing or are in highly competitive environments—find it hard to be overly selective when recruiting. Such organizations might instead need to cultivate value alignment through socialization. Furthermore, many organizations use formal performance assessments to promote and disseminate values, making adherence to values such as teamwork, integrity, and a positive attitude a central criterion (WorldatWork, 2019). We examine the effects of a formal values-based PA system on workers' adherence to core values and explore conditions that support or hinder these effects.

Values-based PA systems can convey and reinforce core values by offering guidelines and effectively promoting behaviors reflecting those values. By integrating inspirational values and guidelines for their implementation, these systems can stimulate learning and self-reflection among both raters and ratees, making assessments more meaningful. However, employees may *not respond* to such systems if they are not linked to explicit incentives, if employees don't envision a future with the company, or if they heavily discount their expected future payoffs (Baker et al., 1988; Holmstrom & Milgrom, 1991).

Moreover, employees might react *negatively* to such systems. For instance, prior experimental studies have found that articulating values in the presence of explicit incentives may be perceived as a constraint, leading to decreases rather than increases in employee productivity (Kachelmeier, Thornock, & Williamson, 2016; Akinyele et al., 2020). Cha and Edmondson (2006) develop a theoretical model whereby the dissemination of values can lead employees to attribute leader behaviors perceived to be incongruent with those values to leaders' hypocrisy (i.e., what the authors call "hypocrisy attribution dynamic"). The authors posit that this, in turn, leads to disillusionment and erosion of trust, which is detrimental to performance. Thus, whether formalizing core values through a values-based PA system on its own (i.e., not

tied to explicit incentives) will lead to intended behaviors and outcomes remains an empirical question. Nonetheless, given that companies implement values-based PA systems intending a positive impact, we formalize the following hypothesis:

Hypothesis 1 (H1): The introduction of a values-based PA system will positively affect behaviors and outcomes consistent with corporate values.

To acknowledge the possibility that variation in organizational contexts (e.g., across business units) might influence the effectiveness of a values-based PA system, in line with existing theories and anecdotal evidence, we pose two research questions, asking which factors may help and which may hinder the efficacy of a values-based PA system.

Research Question 1 (RQ1): What factors <u>enhance</u> the effectiveness of a values-based PA system, positively affecting behaviors and outcomes consistent with corporate values?

A well-designed performance assessment system should enhance organizational alignment through two mechanisms discussed in prior literature: an *evaluative* and a *developmental* mechanism (e.g., Prince & Lawler, 1986; Boswell & Boudreau, 2002). We discuss these mechanisms within the context of a formal *values-based* PA system, which can not only communicate the behaviors desired from employees but also enhance the meaningfulness of their work (Rosso, Dekas, & Wrzesniewski, 2010).

Evaluative Mechanism. The evaluative mechanism of a performance assessment typically involves comparing an employee's performance with expectations or with the performance of other employees and may implicitly affect the employee's rewards and career prospects (e.g., through discretionary bonuses, promotions, or firings). We expect the evaluative mechanism of a values-based PA system to motivate behaviors aligned with the core values it promotes. First, employees want to convince their employers that they deserve a promotion and/or higher pay (Gibbons & Murphy, 1992). Second, to the extent the system conveys inspiring core values, the employees' perceived benefits of pursuing a career with the company are likely to increase (Chatman, 1991; O'Reilly, Chatman, & Caldwell, 1991). The implicit incentive associated with a values-based PA system may also motivate the raters—even if they are not themselves

being appraised by the system. As the raters become more aware of the company's values, they engage in what the company considers desirable behaviors to advance their careers within the company. It is, therefore, important to examine when the evaluative mechanism is likely to be more pronounced.

Research suggests that individuals with greater promotion opportunities are more likely to respond to implicit incentives. For example, in a fast-food chain where promotion decisions were informed by nonfinancial metrics, managers in locations with greater promotion opportunities performed better on those nonfinancial metrics than managers in other locations (Campbell, 2008). To the extent that a values-based PA system motivates effort through the evaluative mechanism—either by clarifying how to be promoted or by enhancing the perceived benefit of pursuing a career in a value-driven organization—we would expect that system to be more positively related to behaviors consistent with corporate values in units offering greater promotion opportunities.

Developmental Mechanism. One of the main benefits of a values-based PA system is that "it can call attention to important performance dimensions heretofore neglected by an organization, in the process of conveying organizational values" (London & Beatty, 1993, p. 361). The developmental mechanism of a values-based performance assessment system should work by communicating the values and behaviors employees need in order to perform well. Such a system can also make the work more meaningful—particularly if the corporate values are inspirational—and trigger discussions among employees on how to incorporate such values into their day-to-day work.

The developmental mechanism encourages employees to (a) draw on relevant skills they already have, (b) acquire, develop, and—where applicable—teach others skills that will improve effectiveness, and (c) adopt (and, for leaders, help their teams adopt) attitudes consistent with the promoted values. This mechanism has the potential to offer clarity, serving employees as both guide and inspiration (Kuvaas, 2006; Kim, Chang, & Kim, 2018). Hence, the developmental mechanism of a values-based PA system is likely to help employees prioritize values-aligned behaviors.

Employees' ability to comprehend and implement corporate values may vary with their tenure and experience. Exposure to top management's decision-making processes helps employees learn about such values and the behaviors consistent with them (Van den Steen, 2010a). Observing, for example, how top management makes a difficult choice—such as whether or not to reduce profits in order to address a customer's concern that could affect the company's long-term prospects—makes clear what values are actually prioritized (Simons, 2010). Research has found that more-experienced managers are more likely to lead their units toward successful tradeoffs. In a large distribution firm that introduced a Balanced Scorecard, Griffith and Neely (2009) find that only units with experienced managers could interpret the scorecard measures and improve performance: "When it is necessary for managers and workers to perform a large number of tasks (e.g., the tasks involved in running a retail establishment), it is important that the manager be able to effectively decide where best to put both his and his workers' marginal effort" (p. 52).

This research suggests that unit managers with relevant experience with corporate values may know better how to implement them. They are therefore likely to be best placed to execute the values promoted by a values-based PA system and to not only develop themselves but also train and coach their teams. Such training and coaching is necessary to enhance employees' understanding of the values conveyed by a performance assessment system (London, Smither, & Adsit, 1997). Without clear developmental guidance, employees may continue to focus on activities they already know how to perform rather than those that are related to newly emphasized values for which they may lack skills (Meglino, 1976).

To the extent that a values-based PA system can increase value alignment through a "developmental mechanism," it should be more effective where unit managers have relevant experience in implementing the values in practice. We therefore examine whether the introduction of a values-based PA system is more likely to be positively related to behaviors and outcomes consistent with corporate values in units led by managers with experience putting those values into action who could transfer their expertise to their team.

Research Question 2 (RQ2): What factors <u>undermine</u> a values-based PA system, negatively affecting behaviors and outcomes consistent with corporate values?

Theory and empirical evidence from prior studies suggest that values-based PA systems might backfire, stifling rather than spurring values-aligned behaviors (Cha & Edmonson, 2006). As leaders draw attention to stated corporate values, they can also lead employees to recalibrate their expectations that those leaders will support and themselves live up to those values. This, in turn, opens the door to hypocrisy attribution to the leaders and associated disillusionment with the core values if leadership is *perceived* to act incongruently with them (Cha & Edmonson, 2006; Thompson & Bunderson, 2003) in two ways described in the next two paragraphs.

Risk of Disillusionment with Leaders' Behavior. Leaders promoting organizational values run the risk of being scrutinized for their own behavior in a way in which they may not have been scrutinized before. This may increase the risk that employees now perceive certain behaviors of their leaders to be incongruent with the espoused values, and, in turn, find their leaders to be duplicitous. Under extreme circumstances, this disillusionment may lead employees to reject both the values-based PA system and their leaders (Cha & Edmondson, 2006; Thompson & Bunderson, 2003). Research suggests that perceived violations of leaders' obligations to uphold values can lead to distrust and withdrawal—the very opposite of the values-based PA system's intended effect (Robinson, 1996; Turnley & Feldman, 2000).

Risk of Disillusionment with Leader's Support. In values-driven organizations, disillusionment can also arise from the misalignment of organizational structures and policies with the espoused values (Cha & Edmondson, 2006; Walton, 1980). For example, employees may become disillusioned with the system if they see management choices as limiting rather than supporting their pursuit of the promoted values. We explore whether variation in the provision of resources necessary to pursue core values, such as customer service, gives rise to negative reactions among employees receiving fewer resources than their peers.

Figure 1 summarizes in blue boxes the factors that could enhance or undermine the effect of the values-based PA system on employee alignment with core values.

----- Insert Figure 1 here -----

3. RESEARCH SETTING AND INTERVENTION

Our research site is a mobile phone retail chain in one of India's main cities (hereafter RETAILER). A typical store in this chain has a manager, a cashier, and a team of promoters representing various brands (e.g., Samsung, Nokia, and Vodafone) whose products and/or services are offered. RETAILER sells a wide variety of products and services, including handsets (cellphones), accessories (e.g., headsets, chargers, cases), and service recharges (i.e., prepaid call minutes or data plans). A significant portion of its customers are repeat customers. To emulate the entrepreneurial spirit, sense of ownership, and incentives of local mom-and-pop stores—the chain's main competitors—the store managers and cashiers are compensated mostly with sales commissions. The promoters are paid by the brands they represent; they are not RETAILER's employees, though they participate in some of its sales incentive plans.

RETAILER seeks to differentiate itself from mom-and-pop stores and to build a long-lasting brand by (a) offering a wider selection, (b) bundling products to enhance customers' perceptions of value (i.e., offering custom solutions to fulfill not only the customers' need for smartphones, but also for credit, insurance, complementary accessories, promotional items, talk time, etc.), and (c) providing trustworthy service (e.g., not selling counterfeit products). The managing director communicates this value proposition to the store staff through personal visits to the stores, weekly store manager meetings at headquarters, and communications via email and the company's information system. She also strongly enforces the focus on trustworthy service by penalizing—sometimes firing—workers for unethical behavior such as theft and misleading customers. While relying on the managing director's personal interactions with store teams had fostered a strong company culture, it was not suited to her ambitions for expansion.

The managing director decided to implement a values-based assessment system to formally communicate and promote the organization's core values in the hopes of shaping desired behaviors among the organization's growing number of employees. The performance assessment system disseminating core values required deciding both, *who would be assessed*, and *by whom*. She decided to focus on assessing store managers and to have all store supervisors and store staff members assess the store managers through a 360-degree system. These choices and the logic behind them are explained below.

3.1 The Store Manager Job

As in most retail chains (Arnold, Palmatier, Grewal, & Sharma, 2009), RETAILER's store managers played many roles and were considered key in shaping behavior and driving store success: they led and trained their teams, communicated inventory and staffing needs to headquarters, and were responsible for good, long-term customer relationships. Store managers worked alongside their staff on a day-to-day basis and frequently interacted with—and made sales to—customers. They were accountable for their stores' ongoing success and were expected to model behaviors conducive to the chain's success.

At the time of our study, RETAILER's store managers were incentivized to behave as owners of the stores they managed. In addition to a fixed monthly salary, they could earn generous sales commissions (with higher rates for more profitable items) but were also held fully accountable for missing items and for selling items at a loss. On average, store managers' variable pay was roughly 140% of their salary and the standard deviation of their pay was as large as their salary.[4] High-powered explicit incentives can motivate workers to behave entrepreneurially but can also lead to excessive focus on short-term, individual-wealth–maximizing activities at the expense of values associated with long-term organizational goals (Holmstrom & Milgrom, 1991; Baker, Gibbons, & Murphy, 1994). And indeed, while the high-powered explicit incentives at RETAILER were congruent with the managing director's emphasis on hard work, they had also introduced behaviors (such as deceiving customers and gaming the system) that were detrimental to the organization's long-term success.

3.2 The Values-based 360-degree System

The values-based PA system (i.e., the 360-degree system) was not tied to any new monetary incentives and was designed as a vehicle to define, communicate, and reinforce the organization's core values and associated desired behaviors. The values-based PA system was centered on the store manager, who was expected to act as a conduit for the implementation and upholding of the core values. By using a 360-degree

[4] These calculations are based on data shared by the company relating to store manager compensation in July 2013. The structure of the store managers' compensation at that time was similar to that during our sample period.

system, the managing director aimed to gain a complete picture of how each store manager was living those values, so as to provide feedback and coaching. In typical 360-degree performance assessment systems, raters include the ratee and his or her subordinates, supervisor, and peers, and may also include customers or suppliers (London & Smither, 1995). At RETAILER, the 360-degree surveys were designed to collect information about each store manager from his supervisor (all store managers were male at the time of our study), the store's cashier, brand promoters working at the store, and the store manager himself.

The managing director sought to gain multiple insights from workers that frequently interacted with the store manager for two reasons (a) to obtain a more informed measure of the store manager's compliance with the core values, which was hard to do solely on the basis of objective performance metrics or the assessment of the supervisor alone,[5] and (b) to engage other staff beyond the store manager into learning about the core values of the organization. Since store staff were best placed to observe their manager in day-to-day operations, their assessments could arguably make for a better performance measurement system. As London and Beatty (1993, p. 360) noted, "Subordinates … may have more complete and accurate information about many leadership behaviors than supervisors have." Feedback to the store manager would then be based on the survey results (including insights from open-ended questions), supplemented with customer evaluations gathered through phone surveys administered by the central office.

The system was presented—to raters and ratees—as a tool that would not only help them learn about the core values, but also help them grow and increase their chances of promotion (thus providing implicit incentives). For the brand promoters, a "promotion" meant being hired as a cashier or store manager. Cashiers were often promoted to store manager and store managers could either be reassigned to more profitable stores (where they could earn significantly higher compensation) or be promoted to "store manager coach" to assist nearby stores (while still earning high-powered financial rewards).[6]

[5] Store managers reported to district managers (supervising an average of 6 stores each), who, in turn, reported to the managing director. Most of the district managers undertook other responsibilities at the central office (related to human resource management, supply chain management, etc.) in addition to supervising their stores.

[6] In the month prior to the intervention, the average total pay (i.e., fixed salary plus sales commissions) of a store manager was approximately 160% that of that of the cashier in the same store. This suggests that cashiers experience strong implicit incentives toward promotions and career advancement. We do not have access to compensation data

Management hoped that by (a) formally communicating core values, (b) periodically asking store teams and supervisors to consider those values while assessing their store manager's behavior (or, for the store manager, his own behavior), (c) providing feedback to store managers, and (d) emphasizing the importance of those values for promotion, they could lead the teams to internalize the values and exhibit value-aligned behaviors. In turn, if store teams could be trusted to behave in the interest of the company's long-term success, the managing director could delegate more responsibilities to them and rely less on direct monitoring. This would lead to greater job satisfaction and motivation for store staff (Van den Steen, 2010b) and better alignment with core company values.

The managing director consulted with our research team about how to design the 360-degree PA system and decided to pilot it in approximately half the stores. She agreed to share company data with our research team to enable us to empirically examine the effects of her intervention.[7] While she was enthusiastic, it was unclear whether store staff members, who up to that point had been focused on monetary rewards, would respond to an intervention untethered from explicit incentives.

The 360-degree surveys asked each respondent (the store manager, store team members, and the manager's supervisor) to assess the manager on behaviors organized according to RETAILER's four core values: (1) "We gain control of our own career by working hard every day and reaching out for support"; (2) "We give more value"; (3) "We are honest and ethical"; and (4) "We are caring and respectful." Respondents were asked where the store manager was doing a good job and where he could do better. The survey the store managers were asked to complete is shown in Appendix 1; it was modified slightly for the store team members and supervisors, primarily to refer to the store manager rather than oneself.[8]

Store selection was randomized (we advised the company on the randomization procedure and oversaw it). However, stores were grouped in "blocks" if they were close to each other to reduce the risk of

for the promoters, who are paid directly by the brands they represent, so we cannot assess the magnitude of the implicit incentives they face.

[7] This opportunity to collaborate with RETAILER arose due to a preexisting relationship between members of the research team and the managing director, who had met at a retailing conference.

[8] As mentioned, we do not use the survey ratings in our analyses. However, we include the survey instrument in Appendix 1 to show the behaviors the company intended to foster with respect to each core value.

contamination if some stores in that group participated and others did not. If any store in a block was randomly selected to be treated, all stores in the block were then selected.[9] Participants in the treatment stores were advised by the managing director that RETAILER was piloting the system only in certain stores as a first phase of implementing it companywide and asked not to discuss it outside their store team.

The managing director launched the system in an interactive session for store staff members (i.e., store manager, cashier, and promoters) of the treated stores. She held the launch session twice, at the end of March 2015 and at the end of April 2015, to include all the treated stores and to help some stores that experienced difficulties accessing the 360-degree surveys online.[10] The presentation had two parts. First, the managing director formalized the company's vision statement and core values (included in Appendix 3) and introduced the values-based 360-degree system designed to reinforce the importance of putting these values into action. In the second part, the attendees completed the values-based 360-degree survey in a computer lab set up for this purpose. The managing director explained that the survey was to be implemented at the stores attending the launch session in order to gain a comprehensive understanding of the support the store manager provided and his commitment to the core values.[11] Attendees were advised

[9] The randomization procedure worked as follows: Each store was assigned a random number using a random number generator. The stores were then ordered, from the lowest number to the highest number. The store with the lowest number was assigned to the treatment condition first, followed by the store with the second lowest number, and so on. Due to geographical proximity that would have posed an undue risk of contamination between treated and control stores, sixteen of the stores were assigned to blocks (there are five blocks in total), while the others remained unassigned to blocks. If a store assigned to the treatment condition was part of a block, then all stores in that block were also assigned to the treatment condition, irrespective of the random number they had been assigned. Then the assignment to the treatment condition continued with the next lowest number. This continued until half of the stores had been assigned to the treatment condition. As a result of this randomization procedure, there are seven stores assigned to three blocks in the treatment group and nine stores assigned to two blocks in the control group. Specifically, of the 16 treated stores in our final sample (see Section 4.1.1), three were included in one block, four were included in two two-store blocks, and nine were not part of any block. Of the 16 control stores, seven were included in one block, two were included in another block, and seven were not part of any block.

[10] In our empirical analyses, we drop all weeks between the first session and the time when the surveys related to the second session were completed by all store team members (7 weeks in total). However, including the observations between the two launches does not change our general inferences.

[11] As shown in Appendix 1, this survey had 48 items. The number of items on survey instruments for upward feedback (in which subordinates provide feedback to their superior) and 360-degree systems can vary widely. For instance, the survey in Hazucha, Hezlett, and Schneider's (1993) study had 122 items, while the survey in Walker and Smither's (1999) study had 29 items.

that the aggregated survey responses (confidentiality of individual responses was assured) would be used to provide feedback and coaching to store managers.[12]

The feedback sessions were held about three months after the launch. By this time, a couple of managers had moved from treated stores to control stores (we address this further in footnote 25) and some of the feedback sessions that should have taken place had not, which reduced the available sample for the period after the feedback sessions. Further complicating matters, there was some variation in how these feedback sessions were conducted (e.g., whether or not the managing director was present). We therefore end our post-period right before the start of the feedback sessions.[13] This could decrease the power of our tests, but it allows us to examine the motivational anticipatory effects of the values-based 360-degree PA system implementation independent of any feedback effects or the sharing of survey results.

4. RESEARCH DESIGN AND EMPIRICAL TESTS

4.1 Research Design

We test the effects of the intervention using archival store-level measures capturing workers' value-alignment that were readily available or could be easily constructed from the organization's pre-existing information system as proxies for the desired behaviors highlighted by the 360-degree assessment survey. Note these measures are not intended to capture every possible way in which store members could adhere

[12] Store staff completed the survey privately. Store team members missing the session also completed the survey privately, outside the store and away from the store manager. The surveys completed by the store managers and the store team members (cashier and promoters) were very favorable; store managers rated themselves an average of 4.8 (out of 5) across all survey questions, while team members gave them an average of 4.4. This suggests that the self-ratings and upward ratings were lenient, as documented in prior research. The store managers' direct supervisors were less favorable, giving them an average of 3.8.

[13] The absence of feedback in our post-period may affect our results, but research suggests that this is not a significant issue. Smither et al. (1995) found managers who received individualized feedback in an upward feedback system no more likely to improve their performance than managers who did not. (In that setting, managers with fewer than three subordinates received only an aggregate organizational report to protect the subordinates' anonymity.) Using the same managers as Smither et al. (1995) but studying them longer, Reilly, Smither, and Vasilopoulos (1996) also find that the performance improvements they identify over four administrations of the feedback system (mostly concentrated early in the system) are unrelated to the number of times managers actually received individualized feedback (which varied from zero to three). They conclude: "Our results suggest that the continued administration of an upward feedback program can result in sustained change over a fairly long period of time and actually receiving feedback may be less important than the exposure to the valued behaviors" (p. 599).

to RETAILER's core values, but to indicate changes in behaviors expected by the company following the implementation of the values-based system. The chosen measures are (1) related to the core values emphasized by the values-based 360-degree system, (2) responsive to workers' actions in the short-term, (3) more objectively measured than those used in prior studies, and (4) less subject to leniency bias than survey ratings. We do not use the results of the 360-degree surveys themselves (i.e., those completed by the store manager, their store team, and their supervisor) as these surveys were conducted only once during our sample period. The availability of value-alignment measures derived from archival data extracted from the company's information system for a treatment and a control group allows us to better assess causal relations relative to prior studies, which have relied predominantly on initial and subsequent performance assessments under the system itself, without reference to a control group.

Our empirical analyses, described in the remainder of this section, examine the value-alignment effects of the implementation of a values-based PA system.

4.1.1 Data and Measures Analyzed

We examine how treated stores performed vis-à-vis control stores on archival measures capturing behaviors related to the four core values assessed through the values-based PA system.[14] For the first core value, "*We gain control of our own career by working hard every day and reaching out for support,"* we measure performance using *Ln(Sales),* the natural logarithm of store-level weekly sales in rupees.[15] This variable captures an output measure that is sensitive to changes in effort by the sales team—that is, store teams that work harder tend to sell more. In addition, stores that reach out to headquarters for additional resources or support, or other stores for inventory needed, etc., are also likely to increase their sales. For the second core value, "*We give more value,*" we use two store-week-level metrics capturing aspects of the service customers receive at the store. These are the net promoter score (*NPS*), which the company

[14] Recall that we examine the initial implementation of a values-based 360-degree PA system, where the implementation included both communicating the company's vision and values and formalizing them through a 360-degree system.
[15] The weekly sales amounts are net of customer returns.

calculated based on customer surveys,[16] and a measure constructed from data recorded in the company's information system capturing the percentage of invoices that included bundles of products and/or services (*% Invoices with Bundles*), reflecting the sales associates' selling of complete solutions to customers.[17] We measure value-alignment on the third core value, "*We are honest and ethical,*" with an indicator identifying the incidence of unfavorable outcomes of periodic random audits (*Failed Audit*). The company audits each store approximately once a month to assess the correspondence between (a) inventory and cash represented in the local ledgers and (b) their physical presence in the store.[18] The results of these audits are recorded in the company's information system. We capture behaviors linked to the fourth core value, "*We are caring and respectful*" with two items included in the customers' survey, namely the perception of the sales associates' politeness (*Politeness*) and helpfulness (*Helpfulness*).[19] Sales performance was already rewarded with monetary rewards in the pre-existing incentive system, and, in some cases, store teams were financially penalized for large cash or inventory shortages identified during the audits. None of the other measures were explicitly linked to monetary incentives nor formally tracked/communicated to store teams.

The measures employed as dependent variables in our analyses capture a subset rather than the entire array of value-consistent behaviors intended to be fostered by the values-based PA system. However, despite being partial indicators, these metrics were endorsed by the managing director as appropriate proxies, effectively capturing observable behaviors aligned with each of the values that the performance

[16] Three weeks before the formal launch of the values-based PA system, company staff had begun collecting customer survey data across all stores (regardless of whether they were control or treatment stores), randomly selecting one or more customers per store each day for a telephone survey and continued this during our sample period. Selected customers were asked about aspects of the service they received (such as the staff's knowledge, helpfulness, and politeness), their overall satisfaction with the service, and how likely they were (on a scale of 0 to 10) to recommend RETAILER to a friend. From this last question, a store's Net Promoter Score (hereafter NPS) was calculated as the percentage of "promoters" (respondents giving a 9 or 10) minus the percentage of "detractors" (respondents giving a score of 0 through 6) - see www.netpromoter.com. The NPS is a popular metric for customer experience. The customer survey is shown in Appendix 2.

[17] As mentioned, offering complete solutions was a competitive advantage for RETAILER over their competitors. Additionally, often items such as headsets and other accessories were included in the bundle free of charge.

[18] Note that there is significant variation in the reasons why an audit could be flagged as a "failed audit." Any inconsistency, large or small, between system data and physically observed items (e.g., cash or inventory in the store) could trigger a "failed audit" flag. In untabulated tests, we find no association between the number of audits and the incidence of failed audits at a store.

[19] Specifically, the customer survey asked, "Were the people who served you polite?" (*Politeness*) and "Did the people who served you go out of their way to help you? (*Helpfulness*). See Appendix 2 and 4 for more details.

assessment system aimed to emphasize. Appendix 4 lists the archival value-alignment measures that we use and their corresponding core values.

Next, we describe measures for the factors that could either enhance or undermine the effectiveness of the values-based PA system. Figure 1 not only summarizes the theoretical constructs associated with these factors (highlighted in blue boxes), but also describes the measures that we employed as proxies for these constructs (depicted in white boxes).

To analyze conditions influencing the evaluative and developmental mechanisms potentially enhancing the effectiveness of the values-based PA system, we define two variables. The variable *Promotion Opportunities* as (# stores within a 1-mile radius) / (# store team members within a 1-mile radius)[20,21] is expected to amplify the evaluative effect of the values-based PA system, increasing the perceived consequences of such evaluation. The second variable captures the store manager's experience with the company core values and is expected to enhance the developmental effect of the system on the store managers' team. We construct an indicator variable *Experienced Manager* equal to one if the store manager has been in the role for at least six months at the time of the introduction of the values-based PA system, and zero otherwise. In a retail chain setting, the experience relevant to a values-based PA system is likely to be more developed amongst store managers who have had prolonged exposure to the company's core values by interacting with top management over a reasonable period of time, leading to greater understanding both of corporate values and of the behaviors needed to incorporate those values into their

[20] The number of store team members competing for promotions is calculated as the median number of store team members working at store *i* in the pre-period, plus the median number of store team members working at stores within a one-mile radius from store *i* in the pre-period, minus one (because team members do not compete with themselves). The number of stores in the one-mile radius includes store *i* to account for the promoters' opportunity to become a cashier or store manager, the promoters' or cashier's opportunity to be promoted to store manager and the opportunity of the store manager to be promoted to store manager coach. We define promotion opportunities within one-mile radius since promotions commonly occur within the same area. Our definition of promotion opportunities is similar to definitions used in prior research. For instance, Campbell (2008) uses a similar indicator of promotion opportunities by identifying market areas where (a) the likelihood of higher-level positions becoming available in the future is high (due to higher concentration of business units), and (b) the number of other candidates for that position is low.

[21] We construct this measure based on the store conditions observed in the three months preceding the implementation of the values-based PA system to ensure that it is relevant to the potential effects of the intervention. We apply the same logic and measure other sources of influence on the system's effectiveness (*Low DSI* and *Late Payment*) over the three months preceding the intervention to ensure that the events associated with them were recent enough in the eyes of the workers to have an influence at the time of the intervention.

day-to-day work. In the setting we study, store managers meet weekly as a group with the managing director. This practice allows them to rapidly increase their understanding of the core values by being directly exposed to the managing director, who models and reinforces the core values through her behaviors, as well as having the opportunity to learn from fellow store managers.[22] The managing director considered a six-month period to be long enough for store managers to learn how to behave according to values and to acquire the ability to teach such values to their store teams.

The variables we use to examine conditions potentially undermining the values-based PA system are informed both by the literature reviewed earlier and insights from interviews at the research site. In late May and early June 2015 (that is, a little over a month after the system launched), the managing director tasked an assistant to conduct follow-up interviews. These were aimed at assessing the sentiment of the store staff members with respect to the values-based PA system (see Appendix 5).[23] The results, shared with us by the managing director, along with our literature review, led us to identify two factors—delays in compensation payments and shortage of inventory—that may have negatively affected the intervention due to disillusionment with the leadership.

While the managing director was generally viewed positively and considered to be supportive, some interviewees voiced complaints in the follow-up interviews conducted by the assistant. We identified two sources of discrepancy between the espoused core values and the practices and policies enacted by management which were likely to lead to disillusionment with the core values and the PA system. These pertained to compensation not being paid on time (which seems contrary to the leadership's alignment with the core value of being honest and trustworthy) and not having enough inventory on hand to satisfy the requests of customers (which appears to be inconsistent with leadership support for the principle of giving more value).

[22] To avoid contamination, the managing director did not explicitly discuss the core values or the values-based PA system during the time of the intervention.
[23] In total, 14 people were interviewed. Therefore, the sample size is too small for us to use the content of the interviews as data in our empirical analyses.

Occasionally, management retained cash payments related to store staff members' compensation for a few days for different reasons (e.g., as it tried to resolve issues such as stock missing from those particular stores). We capture this using *Late Payment*, an indicator variable equal to 1 if the store manager of the store was paid after half of the store managers working in close stores (i.e., stores within a 1-mile radius) at least one time during the last three months included in the pre-period, and 0 otherwise. All variables used in our empirical tests are defined in Appendix 4.

Margins for mobile phone sales are generally small and procurement of inventory requires significant investments in working capital; the devices are largely pre-paid by the retailer and risk obsolescence and damage if kept in-store too long. As a result, management faces a constant trade-off between minimizing the stock on hand and risking that the stores may not always have what the customer wants or enough items to sell bundled solutions. DSI (days of sales in inventory) is calculated as the ratio of the average inventory balance and the average cost of goods sold per day. A lower DSI indicates lower availability of inventory on hand. We measure average DSI for each store over the last three months in the pre-period.[24] We then partition the distribution of pre-period average DSI into quintiles, given that we identify a discontinuity in DSI at the lowest quintile. We capture low inventory with the indicator variable *Low DSI*, which is equal to 1 if the store was in the lowest quintile of the distribution based on pre-period DSI, and 0 otherwise.

Our sample includes weekly observations for 32 stores (16 each in the treatment and control groups) spanning 29 weeks, of which 18 are prior to the intervention (pre-period) and 11 are subsequent to the week when the store staff members of all the treated stores completed the 360-degree surveys (post-period).[25] We

[24] Our decision to use pre-period *Low DSI* is informed by the possibility that changes in sales performance subsequent to the implementation of the values-based PA system might impact the DSI measure even in the absence of changes in inventory management policy.

[25] Our initial sample includes 39 stores. Of these, we drop seven from our analyses. We drop two stores that were supposed to be included in the treatment group because their store managers did not complete a 360-degree survey (in the first case, the store manager had been temporarily reassigned to a control store in April 2015; in the second case, the store manager was absent at the time of the intervention). We drop another treatment store because it opened immediately before the sample period and closed immediately after it. We drop two further stores because the store team participated only in the core values session and not in the 360-degree survey. We also drop one control store that was contaminated by the reallocation of a treated store manager immediately after the intervention. Lastly, we drop one store where the team identified the cashier as the (acting) store manager in their 360-degree surveys, but the cashier himself did not self-identify as store manager in his survey."

drop any post-period observations subsequent to the departure of a treated store manager if there was no replacement during our sample period and any post-period observation in which a control store is contaminated by a treated store manager being reassigned to it. Our final sample includes 918 store-week observations. We have fewer observations for our *NPS, Politeness,* and *Helpfulness* measures, since the company began administering customer surveys only three weeks before formally launching the values-based PA system, as well as for the *Failed Audit* variable, since stores are randomly audited on average once a month, not all stores are audited each month, and we drop any observations pertaining to stores that do not have at least one audit in both the pre- and the post-period. Consequently, we have lower power to detect intervention effects on these measures.

Table 1, Panel A summarizes descriptive statistics for the variables of interest. Sales exhibit significant variation across store-weeks. Despite some instances of a perfect score (100%), *NPS* is at or below 33% for half of the sample.[26] Measures of customers' perceptions related to the service received (i.e., *Helpfulness* and *Politeness*) span across a 1-5 Likert scale, and present, on average, opportunities for improvement. The *% Invoices with Bundles* is below 35% for three-quarters of the sample, suggesting opportunities to offer greater customer value by providing more comprehensive product and service solutions. *Failed Audit* is equal to 1 for more than half of the sample (this is not surprising given cash and inventory shortages are frequently observed in the retail sector (Chen & Sandino, 2012)). With few exceptions (only 5.6% of the store-weeks included in our sample), stores record sales every day of the week each week.[27] About 76% of the store managers had been in the role for at least six months at the time of the intervention. Assuming similar skills and potential among store team members, the average store team member has roughly a 19% probability (varying from less than 10% to 50% depending on the market) of being selected for a new store

[26] The average NPS (34%) is comparable to the average NPS scores of 27% in India and 49% in the United States (Kelly, 2019). In the analysis of variables obtained from the customer surveys (i.e., *NPS*, *Politeness*, and *Helpfulness*), we eliminate two stores for which we have only post-period observations.

[27] Note that the last week of 2014 comprises 8 days in the statistical software package that we use, hence the maximum of 8 for this variable.

position.[28] In the last three months of the pre-intervention period, stores held an average of 25.9 days of sales in inventory and about 13% of store managers had received their compensation late at least once.

Table 1, Panel B reports tests of covariate balance between control and treatment stores *in the pre-period.* Despite random assignment of treatment, we find that treatment stores had higher sales on average in the pre-period relative to control stores, and a lower incidence of failed audits on average. We find no other significant differences. In all our analyses we include store fixed effects to control for time-invariant store characteristics (which address differences in the stores' average pre-period sales and audit results). Given our relatively small sample and the grouping of some stores in blocks it is not surprising that the treatment and control stores differed significantly on two of the many dimensions we examine.[29]

----- Insert Table 1 here -----

Table 2 reports the Pearson correlation coefficients for our variables of interest. Interestingly, sales are not significantly correlated with our measures associated with the quality of customer service, which is in line with the managing director's concern that sales may be driven by aggressive sales tactics rather than customer service in some cases. As expected, the measures extracted from the customer survey data provided by the company are internally consistent, with positive correlations between all combinations of *NPS*, *Helpfulness*, and *Politeness*, except one (*NPS* and *Helpfulness* are not correlated). Sales are positively correlated with providing complete solutions *(% Invoices with Bundles*) and negatively correlated with the availability of inventory on hand (*Average DSI in the Pre-Period*). Stores with greater opportunities for promotion (*Promotion Opportunities*) have lower sales but a higher *NPS*, more frequent store manager changes and less experienced managers. Our overall assessment of the correlations among our predictors is that the risk of collinear relations in the definition of our statistical model is not material.[30]

[28] This is just a high-level interpretation considering the fact that our measure captures multiple types of promotions for which different sets of store staff members could be eligible.
[29] Table 1, Panel B uses all available pre-period observations, except for the variables capturing influencing factors where we use one observation per store as these variables are time-invariant. In untabulated analyses, we adopt an alternative approach where we use only one observation per store, which captures the average value over the pre-period. With this approach, we find no significant pre-period differences between treatment and control stores.
[30] We corroborate our conclusion by running tests of collinearity in the estimation of the coefficients in our model.

----- Insert Table 2 here -----

4.1.2 Analysis of the Average Effects of the Intervention (H1)

We begin by examining the average effects of the intervention on behaviors aligned with the company's four core values. Hypothesis 1 predicts that the introduction of a values-based PA system will positively affect behaviors and outcomes consistent with corporate values. We test our prediction by estimating the following difference-in-differences statistical model:

$$\begin{aligned} Value\ Alignment_{i,t} = {} & \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Store\ Manager\ Change_{i,t} \\ & + \beta_3 Sales\ Days_{i,t} + \beta_n (Store\ Fixed\ Effects) \\ & + \gamma_t (Week\ Fixed\ Effects) + \varepsilon \end{aligned} \quad (1)$$

where the dependent variable *Value Alignment* is substituted by each of the dependent variables described earlier; the indicator variable *Post* equals 1 if week *t* is after the implementation of the values-based PA system, and 0 otherwise; the indicator variable *Treatment* equals 1 if the system is implemented in store *i*, and 0 otherwise;[31] the indicator variable *Store Manager Change* equals 1 if the manager of store *i* was different in week *t* than at the time of the intervention, and 0 otherwise;[32] and *Sales Days* is the number of days in the week where the store had at least one sales transaction. To account for unobservable store-specific factors, all our estimations include store fixed effects. All our estimations also include week fixed effects to account for the associated unobservable heterogeneity. Since we observe repeated measures of value alignment for each store in our sample, standard errors are clustered by store in all our estimations.

4.1.3 Analysis of Factors Influencing the Effect of the Intervention (RQ1 and RQ2)

We explore factors that are likely to influence the response to the introduction of a values-based PA system. In particular, we predict that the availability of greater opportunities for promotion will accentuate the *evaluative* mechanism of a values-based PA system, thus leading to greater value-alignment improvements in stores with greater promotion prospects. To the extent store members saw the values-

[31] The main effect of *Treatment* does not appear in Equation (1) since it is absorbed by the store fixed effects, and the main effect of *Post* does not appear since it is absorbed by the week fixed effects.
[32] This variable is set to zero for all stores in all pre-intervention period weeks.

based PA system as an evaluative mechanism, the system could give rise to stronger implicit incentives in locations with greater opportunities for promotion.

We also expect the effectiveness of the values-based PA system to depend on the store manager's experience with respect to executing the core values, derived from working with top management. To the extent the effectiveness of the intervention depended on the store managers' ability to coach, guide, and develop their teams, we expect the store managers' experience working with top management to be a driver of the system's success.

Certain factors may also undermine the values-based PA system. To the extent that store team members felt that management acted contrary to the values or perceived that they received insufficient support to implement the core values, the effectiveness of the system may have been hindered. This could arise at stores where there is a history of late payments to the store manager or at stores that receive limited resources (i.e., low inventory).

To examine whether these four factors influenced the effect that the values-based PA system had on the outcomes analyzed, we estimate the following model:

$$\begin{aligned} Value\ Alignment_{i,t} = {} & \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Post_t * Factor_i \\ & + \beta_3 Post_t * Treatment_i * Factor_i + \beta_4 Store\ Manager\ Change_{i,t} \\ & + \beta_5 Sales\ Days_{i,t} + \beta_n (Store\ Fixed\ Effects) \\ & + \gamma_t (Week\ Fixed\ Effects) + \varepsilon \end{aligned} \quad (2),$$

where *Factor* is substituted with *Promotion Opportunities*, *Experienced Manager*, *Late Payment, or Low DSI*. Because the variables capturing the factors that we predict will influence our relations are defined as time-invariant store-level characteristics and all our estimations include store fixed effects, *Treatment, Factor,* and the interaction between *Treatment* and *Factor* are subsumed.

4.2 Results of Statistical Tests

4.2.1 Multivariate Tests of the Average Effects of the Values-based PA System (H1)

We perform separate OLS estimations of Equation (1) for each of our six dependent variables of interest. Since Equation (1) is specified using a difference-in-differences approach, our primary coefficient

of interest is β_1, which is associated with the interaction term (*Post*Treatment*). We find no significant differences between treated and control stores in the post-implementation period with respect to any of our dependent variables (see Table 3).[33] Thus, we find no empirical support for Hypothesis 1.

----- Insert Table 3 here -----

However, descriptive evidence from follow-up interviews that the company conducted with a number of store staff members and shared with us indicate that the store managers and their team members generally viewed both the session launching the values-based PA system and the completion of the related 360-degree surveys positively. They expected the system to yield several beneficial effects including enhancing transparency, motivating workers, improving the store environment and teamwork, facilitating support from senior management, and developing store team members. They understood that the system involved obtaining information about the store manager's behavior and some respondents mentioned that the purpose of the system was to motivate the store manager and/or help him to grow/improve.

The disconnect between the overall positive reception of the intervention and the lack of average effects reinforces the relevance of examining factors that may influence the effectiveness of the introduction of a values-based PA system.

4.2.2 Multivariate Tests of the Evaluative Mechanism due to Promotion Opportunities (RQ1)

Table 4 reports our OLS estimation results of Equation (2) where *Factor* is substituted by *Promotion Opportunities*. The coefficient associated with the interaction term (β_3) is positive and significant for three of our dependent variables: *Ln(Sales)*, *NPS,* and *% Invoices with Bundles*. Based on these results, a one standard deviation increase in *Promotion Opportunities* (standard deviation=0.120) is associated with marginal increases of 52.7% in sales (p-value<0.01), 20 percentage points in the *NPS* score (p-value=0.03), and 5.2 points in the percentage of invoices including bundles (p-value=0.09), for treated stores relative to

[33] Directionally, the effect of the intervention is in the expected direction for five of the dependent variables, while it is in the direction opposite to expectations for one of the dependent variables (*Politeness*). Further, we note that although the effect on sales is statistically insignificant at conventional levels, it seems to be substantial, representing a 22% increase in sales following the intervention (calculated as exp^0.201-1), and the effect is close to being statistically significant (p-value=0.13).

control stores, following the introduction of the values-based PA system.[34] In other words, the values-based PA system was associated with more favorable changes in sales, net promoter score, and cross-selling or bundling, the greater the stores' opportunities for promotion.[35]

These results are consistent with insights from a conversation that we had with the managing director where she highlighted that some brand promoters had mentioned that they were especially interested in the intervention because it gave them guidance on how to work towards being promoted.

Taken together, our analyses suggest that the values-based PA system was most effective when the *evaluative mechanism* of the system was more pronounced. Specifically, *Promotion Opportunities* influenced the effect of introducing the values-based PA system on value-alignment dimensions related to two of the four core values the company was introducing (related to working hard and giving more value to customers). These dimensions included outcomes both related and unrelated to pre-existing explicit incentives, suggesting that the system motivated the store staff to exert effort both on short-term and long-term objectives inherent in the company's core values where the opportunities for promotion were higher.

----- Insert Table 4 here -----

4.2.3 Multivariate Tests of the Developmental Mechanism due to Store Manager Experience (RQ1)

Results of an OLS estimation of Equation (2) where *Factor* is substituted by *Experienced Manager* are reported in Table 5. We find significant interaction effects, whereby the incremental effect of having an experienced manager corresponds to a marginal increase of 67.4% in sales (p-value<0.01), and an increase in the *Politeness* rating of 0.84 points (p-value=0.01), and in the *Helpfulness* rating of 1.22 points (p-value<0.01), both of which are measured on a Likert scale from 1 to 5 points with 1 corresponding to the lowest levels and 5 to the highest levels, for treated stores relative to control stores, following the

[34] For ease of interpretation of the estimated coefficients, we have mean-centered the variable *Promotion Opportunities.* When the DV is *Ln(Sales)*, the coefficient for the triple interaction (*Post*Treatment*Promotion Opportunities*) = 3.527. Therefore, the marginal effect on sales for treated stores relative to control stores in the post period associated with one standard deviation for *Promotion Opportunities* is calculated as follows: e^(0.120*3.477)-1 = 0.5269. Similarly, the marginal effects for the *Post*Treatment*Promotion Opportunities* interaction on NPS and the % of Invoices with bundles are (0.120*1.670=20%) and (0.120*0.435=5.2 points) respectively.

[35] We subject all of our statistical analyses to a battery of robustness tests, which we summarize in Appendix 6.

introduction of the values-based PA system.[36] This suggests that, having a store manager with at least six months of experience led to significant marginal intervention effects both on a measure that was previously explicitly incentivized as well as measures that were not-explicitly incentivized, representing two of the four core values promoted by the system related to working hard and being caring and respectful to others. The fact that the store manager's experience enhanced the effectiveness of the values-based PA system on a number of our dependent variables is consistent with the idea that workers were able to acquire or apply skills both to pursue new behaviors and to successfully balance the pursuit of short-term and long-term value-alignment dimensions.

These results are also consistent with a conversation with the managing director, who noted that, following the intervention, several workers approached her to ask how they could create respectful, long-term relationships with their customers. In her words: "They are asking, 'How do we live [the core values] daily? We have too many customers that walk into our stores. So how do we build a relationship with our customers? How would you do it?'" There is considerable scope for long-term relationships with customers in this setting. For example, customers routinely purchase prepaid minutes for their phones. More than a quarter of customers that completed the customer surveys during our sample period were repeat customers.

----- Insert Table 5 here -----

4.2.4 Multivariate Tests of the Risk of Disillusionment with Leaders' Behaviors (RQ2)

Table 6 reports our OLS estimation results of Equation (2) where *Factor* is substituted by the *Late Payments* indicator. We find a negative incremental post-period effect in treatment stores relative to control stores when the store manager was paid late in the pre-period. This incremental negative reaction to the

[36] In untabulated analyses, we replicated Table 5 using an alternative version of *Experienced Manager* equal to 1 if the store manager has more than one year of experience at the store and zero otherwise. We find positive and significant results for the *Post*Treatment*Experienced Manager* interaction when the dependent variables are *Politeness* (p-value = 0.073) and *Helpfulness* (p-value = 0.091), although the statistical significance of their coefficients is lower than in Table 5. These results suggest that the influence of having an experienced manager on the effectiveness of the intervention is better identified by using a cutoff at six months than at one year. This is consistent with the managing director's suggestion that the store managers gained experience quickly, and that the pace of their learning diminished after being a few months in the job.

introduction of the values-based PA system is reflected in *NPS* being 53.2 percentage points lower (p-value<0.01).[37]

----- Insert Tables 6 and 7 here -----

4.2.4 Multivariate Tests of the Risk of Disillusionment with Leaders' Support (RQ2)

Table 7 reports the estimation results of Equation (2) where *Factor* is substituted by the *Low DSI* indicator. We find that low availability of inventory had a negative incremental effect on treated stores relative to control stores following the intervention, reflected in a marginal *reduction* of the percentage of invoices including bundles by about 17.4 percentage points (β_4=-0.174, p-value<0.01). This result is consistent with the intervention backfiring where store teams may have responded negatively to the company's demand to sell complete solutions to customers when they were not equipped with the inventory necessary to achieve this objective.

The results of this study were informative to the managing director, who then took steps to further promote and clarify the company's core values and to improve support, including paying compensation on time and increasing support to store managers through coaching and other systems. Thus, while in the short-term a values-based PA system could have negative effects if it causes disillusionment with management, it could have a net positive effect in the longer term if it compels leadership to adopt behaviors that are more aligned with company values.

5. CONCLUSIONS

This study explores the effects of implementing a values-based PA system in an Indian mobile-phone retail chain. This system was introduced in half of the chain's stores to formalize and disseminate the company's vision and core values. It aimed to introduce developmental and evaluative mechanisms that would motivate workers to balance short-term results with long-term organizational objectives.

[37] We estimate a negative interaction term also when the dependent variable is *Ln(Sales)*, consistent with the system backfiring in these conditions. However, the result is only one-tail significant (β = -0.438; p = 0.12).

We examine the effects of the introduction of the system—using archival company data, captured before and after the intervention and before any feedback was provided—on outcomes that were immediately actionable and largely controllable by workers in their day-to-day work. The intervention yielded interesting results. On average, there were no significant value-alignment effects on observable proxies. However, the system's effectiveness exhibited significant heterogeneity. Specifically, workers' alignment with values improved when the evaluative and/or developmental mechanisms of the values-based PA system were accentuated. First, the system was more effective in stores where store managers and their teams had greater opportunities for promotion. In these stores, the evaluative role of the values-based PA system was enhanced, as workers (including those appraised as well as those who completed surveys to appraise others) had implicit incentives to act in accordance with company values in order to be promoted. Second, the system was more effective where store managers had sufficient experience related to acting on the core values. These results are consistent with the developmental role being enhanced as store managers were more likely to model and promote behaviors aligned with core values among their teams. Conversely, we find that factors potentially driving workers' disillusionment with the values-based PA system, such as perceiving that the leadership was not upholding the core values (as reflected in delayed payments to some store teams), or, that it was not providing adequate support to pursue core values (in this case, inventory to deliver on the promise of giving more value to customers) led to a deterioration of value-aligned behaviors after the intervention. This suggests that perceived inconsistencies between leaders' actions and espoused values could demotivate certain workers and help explain why the overall average results of the system implementation were insignificant.

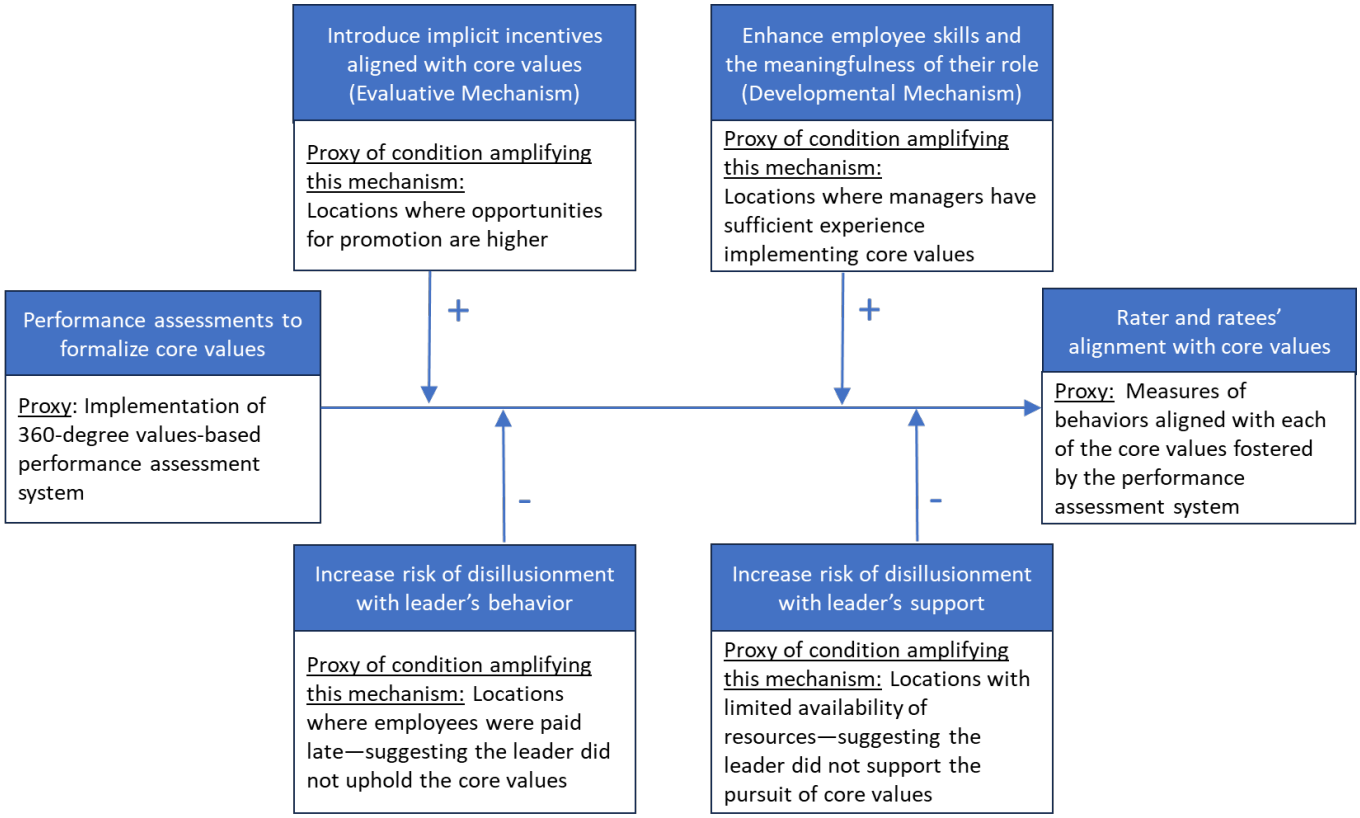
Our findings contribute to an emerging literature on the role of formal management control systems in disseminating corporate values. While most prior studies have focused on formal recruitment systems, we provide novel insights into how and when formal management controls can instill corporate values among workers. Drawing on theory, we highlight the developmental and evaluative mechanisms of performance assessments, particularly investigating their role in aligning employees to core values. And we distinctively

highlight potential demotivating factors that may arise from perceived incongruences between leadership actions and espoused values, which could derail a successful implementation of a values-based PA system.

Our results have implications for companies considering the formalization of core values using an assessment system. However, we acknowledge various limitations. First, the organization we study decided to formalize, and assess its workers' alignment with, core values as part of the same intervention (i.e., the values-based 360-degree assessment system was the vehicle through which the managing director formalized the core values). Future research could explore the effects of a values-based PA system in a setting where the core values were previously formalized. Second, our setting leveraged a 360-degree performance assessment system, motivated by the fact that adherence to core values is difficult to assess unidirectionally. Future studies could explore the effectiveness of values-based assessments performed solely by the workers' supervisors. Third, we acknowledge that the influencing factors that we examine were not randomly assigned to the stores in our treatment and control groups and that there are some correlations between the factors and other variables of interest (both predictors and outcomes). Having said that, we show in our covariate analysis that our randomization strategy resulted in treatment and control groups that had no significant differences among these factors. Furthermore, we explain that our difference-in-differences analysis helps control for any differences in the level of these influencing factors as the stores serve as their own control. Additionally, we mitigate endogeneity concerns by measuring these factors in the pre-intervention period. Fourth, our relatively small sample and potentially incomplete proxies for the company's core values may have limited our ability to detect significant results. However, our archival proxies allow us to conduct difference-in-differences analyses and to use more unbiased measures than the ratings traditionally used to test the effects of performance assessment interventions. Most critically, we find significant results on the analyses that have the greatest potential to be generalizable: conditions derived from theory affecting the effectiveness of a values-based PA system. Lastly, while our study focuses on a single organization, our results can be reasonably generalized according to List's (2020) "SANS" criteria: Regarding the "S" for selection, our study represents a typical brick-and-mortar retail or service chain, though operating in an emerging market, and the subjects did not self-select ensuring unbiased

representation within our setting. Furthermore, our study relies on theoretical arguments that are not limited to a particular industry. Regarding "attrition" and "naturalness" (the "AN" criteria), the treated stores smoothly integrated the system as part of their regular operations, without engaging in attrition. This allows us to capture genuine employee reactions. Regarding the final "S" for "scalability," the program was later rolled out throughout the organization. The managing director applied insights of the study to "correct course" on the factors initially associated with negative reactions. Overall, our focus on a single organization allows us to gain deeper understanding of workers' reactions to a values-based PA system.

Figure 1: Summary of constructs and proxies



Notes: This figure visually depicts the theoretical constructs considered in our study (highlighted in blue boxes) alongside the proxies we use to empirically test those relationships (indicated in the white boxes beneath each blue box). The signs capture whether we expect the condition to enhance (positive sign) or undermine (negative sign) the expected positive effect of a values-based PA system.

Appendix 1: RETAILER's Store Manager Values-Based 360-Degree Survey

Core Value #1: **We Gain Control of Our Own Career By Working Hard Every Day and Reaching Out for Support**
We work hard every day to grow and succeed in life, and to make RETAILER successful. We know that by working hard and honestly, we can earn more and we have the chance to be promoted. We reach out to other stores, head office, brands, managers, and the distribution center (DC), to get stock, to get numbers activated, and to get repair and dead on arrival (DOA) cases resolved. By reaching out we achieve higher sales, make more money and have happier customers who will recommend our stores.
Select a number between 1 and 5 for every question where 1 means Never, 2 means Rarely, 3 means Sometimes, 4 means Very often, and 5 means Always.
1.1 Do you explain to all promoters and the cashier their targets and the reward program?
1.2 Do you give all promoters and the cashier their daily target every morning?
1.3 Do you ask all promoters and the cashier every day about their target achievement till date?
1.4 Do you remain positive about targets even if there have been some bad days?
1.5 Do you try to increase sales by reaching out to customers outside the store (for example, by distributing leaflets, making posters of special offers, or telling promoters to stand at the canopy outside the store)?
1.6 Do you work hard to help everyone achieve targets even if popular models are out of stock?
1.7 Do you tell all promoters and the cashier to sell old and stuck models?
1.8 Do you sell to the customer when the brand's promoter is not present?
1.9 Do you make sure any problems in the store (PC, printer, AC, lights, sign board problems, etc.) get fixed?
1.10 Do you stay late if a customer walks into the store at the time of store closing?
1.11 Do you make efforts to get a new promoter if a current promoter resigns or is absent for a long time?
1.12 Do you help get stock from other stores when needed?
1.13 Do you make all the promoters and the cashier believe that they can have a successful career at RETAILER?
1.13b. Do you believe that you can have a successful career at RETAILER?

Appendix 1: RETAILER's Store Manager Values-Based 360-Degree Survey (Continuation)

Core Value #2: We Give More Value
We give the best combo offers to our customers. For example, our handset plus headphone offer and our handset plus insurance offer are the best value in the market. We constantly try to learn about the products and services we sell so that we can know what options to offer to our customers and answer their questions better. We help customers with their problems in any way we can. We work to build long-term relationships with our customers so that they will visit RETAILER again.
Select a number between 1 and 5 for every question where 1 means Never, 2 means Rarely, 3 means Sometimes, 4 means Very often, and 5 means Always
2.1 Do you know all the DPs, schemes and offers from the different brands (example, Bajaj Finance, EMI scheme, Cash back, PayTM Scheme, etc.)?
2.2 Do you make sure that the price list, posters, and banners in the store are up-to-date?
2.3 Do you know the local market prices?
2.4 Do you teach less experienced team members how to sell profitable bundles to the customer (for example, by giving free gifts, free apps, insurance, unlimited calling, etc. to close sales at a higher price)?
2.5 Do you try hard to close sales against competitors without lowering the price?
2.6 Do you contact the store's customers when the out of stock products become available in the store?
2.7 Do you contact previous customers to tell them about new products?
2.8 Do you make long-term relationships with customers?
2.9 Do you ask all promoters and the cashier to make long-term relationships with customers?
2.10 Do you make accurate commitments to the customer? (For example, you do not promise that a number will be activated in 3 days, or promise that a handset will be repaired within a certain amount of time if you do not know when it will be repaired) In this question, NEVER means you make promises you do not know RETAILER can fulfill, such as those in the examples, to all of your customers SOMETIMES means you make promises you do not know RETAILER can fulfill, such as those in the examples, to about half of your customers ALWAYS means you never make promises you do not know RETAILER can fulfill, such as those in the examples, to any of your customers
2.11 Do you instruct all promoters and the cashier to make only accurate commitments to the customer?
2.12 Do you help customers that have problems? (Some examples of helping are giving your mobile number to the customer, solving activation problems, showing the customer the way to the service center, or even personally going to the service center with the customer, or sending some person from the store to the service center with the customer) In this question, NEVER means you never take any action to help the store's customers that have problems SOMETIMES means you take one or more actions, such as those described in the examples, to help about half of the store's customers that have problems ALWAYS means you take one or more actions, such as those described in the examples, to help all of the store's customers that have problems
2.13 Do you tell everyone in the store to help customers with problems?
2.14 Do you tell everyone in the store to give the same respect to all customers regardless of their purchase amount? (For example, to be equally respectful to a customer wanting a small Rs. 10 recharge and a customer buying an apple phone)
2.15 Do you tell everyone in the store to be respectful to irritated customers?

Appendix 1: RETAILER's Store Manager Values-Based 360-Degree Survey (Continuation)

Core Value #3: We are Honest and Ethical
We are always honest and ethical and we do the right thing at the store every day. We believe that this is the only way to make our store and RETAILER successful.
Select a number between 1 and 5 for every question where 1 means Never, 2 means Rarely, 3 means Sometimes, 4 means Very often, and 5 means Always.
3.1 Are you trustworthy to customers? (Examples of NOT being trustworthy are telling lies about what is being sold to the customer, selling fake products, changing the original batteries of the handset for cheaper batteries, taking the customers' money based on false promises) In this question, NEVER means you take at least one action that is "not trustworthy," such as those described in the examples, with all of the customers you serve SOMETIMES means you take at least one action that is "not trustworthy", such as those described in the examples, with about half of the customers you serve ALWAYS means you never take an action that is "not trustworthy" with any of the customers you serve
3.2 Do you tell all promoters and the cashier to be trustworthy to customers?
3.3 Do you stop wrong activity against the company? (Examples of wrong activities are: stealing, lying, giving unauthorized discounts to friends or family, selling products that are not coming from HO at the store, people making profits for themselves when serving a customer, borrowing store cash or allowing someone to borrow store cash) In this question, NEVER means you never stop wrong activities SOMETIMES means you stop about half of the wrong activities that you notice, such as those described in the examples ALWAYS means you stop all of the wrong activities that you notice, such as those described in the examples
3.4 Do you report wrong activity against the company to HO? Please select "Cannot Answer" if there hasn't been any wrong activity.
3.5 Are you honest at the store? (Being honest means not doing any wrong activity) In this question, NEVER means you do at least one wrong activity, such as those described in the examples, one or more times a day SOMETIMES means you do at least one wrong activity, such as those described in the examples, about once a week ALWAYS means you never do any wrong activity
3.6 Do you tell the cashier and the promoters to be always honest in the store?
3.7 Do you transfer promoters out of your store for personal issues?

Appendix 1: RETAILER's Store Manager Values-Based 360-Degree Survey (Continuation)

Core Value #4: We are Caring and Respectful
We care about and respect each other and our customers. We help each other to grow and be more successful. This is who we are.
Select a number between 1 and 5 for every question where 1 means Never, 2 means Rarely, 3 means Sometimes, 4 means Very often, and 5 means Always.
4.1 Do you give the same respect to customers regardless of caste, religion, gender, or economic status?
4.2 Do you give the same respect to the promoters and the cashier regardless of caste, religion or gender?
4.3 Do you earn the trust of the promoters and the cashier?
4.4 Do you tell all promoters and the cashier when they've done a good job (using words such as Well Done, Good job, Keep it up)?
4.5 Do you tell people at the store to work as a team?
4.6 Do you help solve any fights among the promoters or between the cashier and the promoters at the store?
4.7 Do you understand the personal problems of the cashier and the promoters?
4.8 Do you ask for help from HO when a promoter or the cashier needs it?
4.9 Do you care about the promoters' and the cashier's personal development?
4.10 Do you make efforts to learn about new products and services?
4.11 Do you make efforts to learn from the most experienced promoters?
4.12 Do you make efforts to learn from the WhatsApp group?
4.13 Do you try to learn new things?

Note: We report the entire 360-degree survey to illustrate the behaviors that management expected store managers to exhibit in relation to each of the four core values. This survey was modified slightly for the store members and supervisor, primarily to refer to the store manager rather than oneself. The company administered the 360-degree survey for its own purposes. Management shared the ratings collected using this survey with us. However, in our analyses, we do not use the data collected from the surveys since the 360-degree surveys were administered only once during our sample period. However, the behaviors highlighted by the survey with respect to each of the core values informed our choice of dependent variables used in our statistical tests. We constructed these variables using measures extracted from other sources of data the company shared with us (i.e., company information systems and customer surveys) to capture behaviors aligned with core values, which should have been impacted by the company's pilot implementation of the 360-degree system.

Appendix 2: RETAILER's Customer Satisfaction Survey

Good morning / afternoon / evening,

My name is [name] and I'm calling from RETAILER. You made a purchase from one of our RETAILER stores a few days ago. I'd like to ask you a few questions to get your thoughts and opinions about your visit to our store. This should take about 4 or 5 minutes. Would that be OK?

[Wait for customer's reply.] Thank you.

1. What did you purchase on your last visit to RETAILER? I will read some possible products and services. Please tell me which ones apply. You can also tell me a different product or service. [Select all that apply. You can choose more than one.]
O Handset
O Recharge
O Connection / Data Card
O Accessories
O Insurance
O Other

2a. Was this the first time you purchased from RETAILER?
O Yes
O No

[If Yes, go to question 3.]

2b. How many times have you made a purchase from RETAILER, including your most recent purchase?
O 2
O 3
O 4
O 5 +

Appendix 2: RETAILER's Customer Satisfaction Survey (Continuation)

3. Why did you choose to buy at RETAILER rather than another store? I will read some possible reasons. Please tell me which ones apply. You can also tell me a different reason. [Select all that apply. You can choose more than one.]
O They have a convenient location.
O They were recommended by a friend.
O They had the product that I was looking for in stock.
O They give live demos.
O They provide good value.
O They offered financing options.
O They offered to buy back my old phone.
O I know someone at the store.
O Other

4. We would like you to answer the following questions based on your most recent visit to RETAILER. 4a. Did the selection of products and services meet your expectations? Please choose a number between 1 and 5 where 1 means not at all, 5 means very much so, and 3 means neutral.
O 1 Not at all
O 2
O 3 Neutral
O 4
O 5 Very much so

4b. Were the people who served you polite? Please choose a number between 1 and 5 where 1 means not at all, 5 means very much so, and 3 means neutral.
O 1 Not at all
O 2
O 3 Neutral
O 4
O 5 Very much so

Appendix 2: RETAILER's Customer Satisfaction Survey (Continuation)

4c. Were the people that served you able to explain the features of the products and services in detail? Please choose a number between 1 and 5 where 1 means not at all, 5 means very much so, and 3 means neutral.
O 1 Not at all
O 2
O 3 Neutral
O 4
O 5 Very much so

4d. Did the people that served you go out of their way to help you? Please choose a number between 1 and 5 where 1 means not at all, 5 means very much so, and 3 means neutral.
O 1 Not at all
O 2
O 3 Neutral
O 4
O 5 Very much so

4e. Were the products and services you received worth the price you paid for them? Please choose a number between 1 and 5 where 1 means not at all, 5 means very much so, and 3 means neutral.
O 1 Not at all
O 2
O 3 Neutral
O 4
O 5 Very much so

5. Were you satisfied with your overall experience at RETAILER? Please choose a number between 1 and 5 where 1 means not at all, 5 means very much so, and 3 means neutral.
O 1 Not at all
O 2
O 3 Neutral
O 4
O 5 Very much so

Appendix 2: RETAILER's Customer Satisfaction Survey (Continuation)

6. How likely is it you would recommend RETAILER to a friend? Please note that for this question we will ask you to choose a number between 0 and 10, not 1 and 5. Please choose a number between 0 and 10 where 0 means not at all likely, 10 means extremely likely.
O 0 Not At All Likely
O 1
O 2
O 3
O 4
O 5
O 6
O 7
O 8
O 9
O 10 Extremely Likely

7. Please explain the reason for your answer to the last question about how likely it is that you would recommend RETAILER.

8. Are there any other additional comments you'd like to make?

[Note to Surveyor: If the customer has a complaint/grievance, please enter the details below.]

That completes our survey. Thank you for taking the time to answer our questions!

Appendix 3: Vision and Core Values at RETAILER

Vision: To build <u>together</u> the largest and most successful mobile phone retailer in India, providing our people with maximum opportunities for growth	
Core values	**Measures used by researchers as proxies of alignment with the core values**
We Gain Control of Our Own Career By Working Hard Every Day and Reaching Out for Support	*Ln(Sales)*
We Give More Value	*NPS, % Invoices with Bundles*
We are Honest and Ethical	*Failed Audit*
We are Caring and Respectful	*Politeness, Helpfulness*

Note: The variables *Ln(Sales)*, *% Invoices with Bundles*, and *Failed Audit* were constructed using data extracted from the company's information systems. The variables *NPS*, *Politeness*, and *Helpfulness* were extracted from the company's customer surveys.

Appendix 4: Definition of Variables Used in This Study

Variable	Definition
Dependent Variables	
Ln(Sales)	Natural logarithm of weekly net sales in rupees for store *i* in week *t*.
NPS	Net promoter score for store *i* in week *t*. This variable is constructed based on a question asking customers how likely they are (on a 0 to 10 scale) to recommend RETAILER to a friend. From this question, we calculate the Net Promoter Score as the percentage of "promoters" (respondents giving a 9 or 10 rating) minus the percentage of "detractors" (respondents giving a score of 0 through 6).
Politeness	Average score awarded by surveyed customers to store *i* in week *t* with respect to the question "Were the people who served you polite?" Individual customers were asked to respond using a scale from 1 to 5, where 1 corresponds to "Not at all" and 5 corresponds to "Very much so."
Helpfulness	Average score awarded by surveyed customers to store *i* in week *t* with respect to the question "Did the people who served you go out of their way to help you?" Individual customers were asked to respond using a scale from 1 to 5, where 1 corresponds to "Not at all" and 5 corresponds to "Very much so."
% Invoices with Bundles	Percentage of invoices with bundles of products or services for store *i* in week *t*.
Failed Audit	Indicator variable equal to 1 if store *i* does not pass a random audit examining missing inventory or cash in week *t*, and zero otherwise.
Explanatory Variables	
Post	Indicator variable equal to 1 if the week/month is after the introduction of the values-based 360-degree PA system, and 0 otherwise.
Treatment	Indicator variable equal to 1 if store *i* is in the treatment group, and 0 otherwise.
Store Manager Change	Indicator variable equal to 1 if the store manager at store *i* in week *t* is different than the store manager that was at the store at the time of the intervention, and 0 otherwise.[38] The variable is equal to 0 in the pre-period.
Sales Days	Number of days where store *i* had at least one sales transaction in week *t*.

(Appendix 4 continues on the next page)

[38] Note that the variable is specified at the store/week level. For example, if the store manager changes for just one week and then the resident store manager comes back the week after, then the indicator variable is equal to 1 for the week in which a different store manager was in the store and goes back to 0 when the resident store manager comes back.

Appendix 4 – Continued

Variables Capturing Influencing Factors	
Promotion Opportunities	Variable calculated as (# stores within a 1-mile radius) / (# store staff members competing for a potential promotion within a 1-mile radius). The number of store staff members competing for a promotion is calculated as the median number of staff members working in store *i* plus the median number of staff members working at stores within a mile-radius from store *i,* calculated over the last three months included in the pre-period, minus one (because an employee does not compete with himself). For the purpose of this definition, store staff members include the store manager, the cashier, and the promoters.
Experienced Manager	Indicator variable equal to 1if the store manager has been in the role for at least six months (counted at the end of the last week preceding the post period), and 0 otherwise.
Late Payment	Indicator variable equal to 1 if the store manager at store *i* was paid at least one time after at least half of the stores located in a one-mile radius from store *i* in the last three months in the pre-period, and 0 otherwise.
Low DSI (Pre-Period)	Indicator variable equal to 1if the average value of monthly days of sales in inventory in the last three months of the pre-period for the store was in the bottom quintile, and 0 otherwise.

Appendix 5: RETAILER's Guide for Follow-up Store Staff Interviews

Question Number	Question
Q1	When did you start working at RETAILER?
Q2.a	Is there anything you like about working at RETAILER? If so, what?
Q2.b	Would you recommend a friend or family member to work at RETAILER?
Q3	Is there anything you don't like about working at RETAILER? If so, what?
Q4	What are your career plans for the future?
Q5	Can you tell me what RETAILER's core values are?
Q6	Do you know what each of these core values mean?
Q7.a	What did you like about the session introducing RETAILER's vision?
Q7.b	Do you have any concerns about the session?
Q8	What do you think is the purpose of the new 360-degree feedback system?
Q9	What is your understanding about who will complete the surveys?
Q10	What is your understanding about who will receive feedback?
Q11	How did you feel about completing a survey about your store manager/yourself for the new 360-degree feedback system?
Q12	What impact, if any, do you think the new 360-degree feedback system will have?
Q13	Has the store manager/Have you discussed the new 360-degree feedback system with your store team?
Q14	(Question for store managers only): How did you feel about others completing a survey about you for the new 360-degree feedback system?
Q15	(Question for store managers only): How do you feel about the upcoming feedback sessions where you will see the results of the surveys and discuss the results with your supervisor?

Note: RETAILER's managing director tasked an assistant to conduct a limited set of follow-up interviews to gather information on the workers' perceptions of the values-based PA system. The company agreed to share the content of these interviews with our research team. Insights from these interviews motivated our examination of influencing factors that could have hampered the effectiveness of the system (i.e., insufficient inventory and being paid late). We constructed these variables from data extracted from the company's information systems. We note that we did not use any of the items included in the interview protocol as data in our study.

Appendix 6: Summary of Robustness Tests

	Overall Intervention (Table 3)	*Influence of Promotion Opportunities (Table 4)*	*Influence of Store Manager Experience (Table 5)*	*Influence of Late Payments (Table 6)*	*Influence of Low Availability of Inventory (Table 7)*
Main Tests (Reported in Tables 3-7)	No Significant Effects	Improved *Ln(Sales)*, *NPS*, *% Invoices with Bundles*	Improved *Ln(Sales)*, *Politeness*, *Helpfulness*	Declined *NPS*	Declined *% Invoices with Bundles*
Robustness Tests – Are the Coefficients of the Main Variables of Interest Consistent with the Main Tests (Yes/No/No-opp)[39]					
Excluding completely one control store in which a treated store manager became the store manager starting from July 2015[40]	Yes*	Yes (*Ln(Sales), NPS*) No *(% Invoices with Bundles*)	Yes	Yes	Yes
Excluding completely one control store in which a treated store manager became the store manager starting from June 2015[41]	Yes*	Yes	Yes	Yes	Yes[42]
Excluding completely a store for which we only had two observations in the pre-period for the customer survey-related variables.	Yes*	Yes (*Ln(Sales), NPS*) No *(% Invoices with Bundles*)	Yes	Yes	Yes[43]

(This table continues on the next page)

[39] We focus on the main variables of interest for each of our main tables. For Table 3, this is the coefficient estimated for the interaction term *Post*Treatment*. For tables 4 through 7, the coefficient of interest is the one estimated for the triple interaction term *Post*Treatment*Factor*, where the source of influence is indicated in the header of each column in this Appendix. "Yes" denotes significance at the 10% level or better in the direction reported in the main tables; "No" denotes non-significant results; and "No-opp" denotes significance at the 10% level or better, but in the opposite direction than what is reported in the main tables. "Yes*" indicates that the results are consistent with those reported in Table 3, which were not significant. Any additional results are reported in footnotes.

[40] In our main tests, we had dropped all weeks after the change in the store manager.

[41] In our main tests, we had dropped all weeks after the change in the store manager.

[42] We also find a decline in Failed Audit (β = -0.850, $p < 0.05$)

[43] We also find a decline in Failed Audit (β = -1.006, $p < 0.10$)

(Appendix 6 – cont'd)

	Overall Intervention (Table 3)	*Influence of Promotion Opportunities (Table 4)*	*Influence of Store Manager Experience (Table 5)*	*Influence of Late Payments (Table 6)*	*Influence of Low Availability of Inventory (Table 7)*
Main Tests (Reported in Tables 3-7)	No Significant Effects	Improved *Ln(Sales)*, *NPS*, *% Invoices with Bundles*	Improved *Ln(Sales)*, *Politeness*, *Helpfulness*	Declined *Ln(Sales)*, *NPS*, *Politeness*	Declined *% Invoices with Bundles*
Robustness Tests – Are the Coefficients of the Main Variables of Interest Consistent with the Main Tests (Yes/No/No-opp)?					
Restrict sample to store-weeks with no change in the store manager	Yes*	Yes	Yes[44]	Yes (*NPS* *No(Politeness, Ln(Sales)*)	Yes
Using *Open Days* instead of *Sales Days*	Yes*	Yes	Yes	Yes	Yes[45]
Dropping stores with no close stores	Yes*	Yes (*NPS, Ln(Sales)*), No (*% Invoices with Bundles*)	Yes	Yes (*NPS)* No (*Ln(Sales), Politeness*)	Yes
Removing *Store Manager Change* from the predictors	Yes*	Yes (*Ln(Sales)*, *NPS*) No *(% Invoices with Bundles*)[46]	Yes	Yes (*NPS)* No (*Ln(Sales), Politeness*)	Yes
Clustering standard errors at the level of randomization	Yes*	Yes	Yes	Yes (*NPS*) No (*Ln(Sales), Politeness*)	Yes

(This table continues on the next page)

[44] We also find a decline in Failed Audit (β = -0.953, p < 0.01)
[45] We also find a decline in Failed Audit (β = -0.894, p < 0.10)
[46]Additionally, we find a negative coefficient when the DV is *Failed Audits*, indicating an improvement in the weekly audits' performance.

(Appendix 6 – cont'd)

	Overall Intervention (Table 3)	*Influence of Promotion Opportunities (Table 4)*	*Influence of Store Manager Experience (Table 5)*	*Influence of Late Payments (Table 6)*	*Influence of Low Availability of Inventory (Table 7)*
Main Tests (Reported in Tables 3-7)	No Significant Effects	Improved *Ln(Sales)*, *NPS*, *% Invoices with Bundles*	Improved *Ln(Sales), Politeness*, *Helpfulness*	Declined *Ln(Sales)*, *NPS*, *Politeness*	Declined *% Invoices with Bundles*
Robustness Tests – Are the Coefficients of the Main Variables of Interest Consistent with the Main Tests (Yes/No/No-opp)?					
Estimations using wild bootstrapping to account for the low number of clusters	Yes*	Yes (*Ln(Sales)*, *NPS*) No *(% Invoices with Bundles*)	Yes	Yes (*NPS*) No (*Ln(Sales), Politeness*)	Yes
Multiple hypothesis testing p-values correction (Romano-Wolf stepdown procedure)[47]	Yes*	Yes	Yes	Yes (*NPS*) No (*Ln(Sales), Politeness*)	Yes

[47] The multiple hypothesis tests pertain to 5 of our 6 dependent variables. We imposed sample restrictions for our analyses when the dependent variable is *Failed Audit* (i.e., certain stores were excluded) and this was not compatible with the multiple hypothesis testing.

References:

Akinyele, K. O., V. Arnold, and S. G. Sutton. 2020. Wording and Saliency Matter: The Impact of Incentive System and Organizational Value Statement on Employees' Performance. *Behavioral Research in Accounting* 32(1): 101-118.

Arnold, T. J., R. W. Palmatier, D. Grewal, and A. Sharma. 2009. Understanding Retail Managers' Role in the Sales of Products and Services. *Journal of Retailing* 85(2): 129-144.

Baker, G. P., R. Gibbons, and K. J. Murphy. 1994. Subjective Performance Measures in Optimal Incentive Contracts. *The Quarterly Journal of Economics* 109 (4):1125-1156.

Baker, G. P., M. C. Jensen, and K. J. Murphy. 1988. Compensation and Incentives: Practice vs. Theory. *Journal of Finance* 43(3): 593-616.

Bernstein, E. S., and S. Li. 2017. Seeing Where you Stand: From Performance Feedback to Performance Transparency. *Academy of Management Proceedings*.

Blanes i Vidal, J., and M. Nossol. 2011.Tournaments Without Prizes: Evidence from Personnel Records. *Management Science* 57(10): 1721-1736.

Boswell, W. R., and J. W. Boudreau. 2002. Separating the Developmental and Evaluative Performance Appraisal Uses. *Journal of Business and Psychology* 16(3): 391-412.

Bradler, C. R. Dur, S. Neckermann, and A. Non. 2016. Employee Recognition and Performance: A Field Experiment. *Management Science* 62(11): 3085-3099.

Bryan, C. J., E. Tipton, and D. S. Yeager. 2021. Behavioral Science is Unlikely to Change the World Without a Heterogeneity Revolution. *Nature Human Behavior* 5: 980-989.

Cai, W. 2023. Formalizing the Informal: Adopting a Formal Culture-fit Measurement System in the Employee Selection Process. *The Accounting Review* 98(3): 47-70.

Campbell, D. 2008. Nonfinancial Performance Measures and Promotion-based Incentives. *Journal of Accounting Research* 46(2): 297-332.

Campbell, D. 2012. Employee selection as a control system. *Journal of Accounting Research* 50 (4): 931–966.

Campbell, D. and T. Sandino. 2019. Sustaining Corporate Culture in a Growing Organization. Harvard Business School Case Note No. 9-119-109.

Casas-Arce, P., C. Deller, F. A. Martinez-Jerez, and J. M. Narciso. 2021. Knowing That You Know: Incentive Effects of Relative Performance Disclosure. Forthcoming at *Review of Accounting Studies*.

Castillo, J. 2002. A Note on the Concept of Tacit Knowledge. *Journal of Management Inquiry* 11(1): 46-57.

Cha, S. E., & Edmondson, A. C. (2006). When values backfire: Leadership, attribution, and disenchantment in a values-driven organization. *The Leadership Quarterly*, *17*(1), 57–78.

Chatman, J.A. 1991. Matching people and organizations: selection and socialization in public accounting firms. *Administrative Science Quarterly* 36 (3): 459-484.

Chen, C. X., and T. Sandino. 2012. Can Wages Buy Honesty? The Relationship Between Relative Wages and Employee Theft. *Journal of Accounting Research*, 50(4): 967-1000.

Gallani, S. 2022. Conduit Incentives: Eliciting Cooperation from Workers Outside of Managers' Control. *The Accounting Review* 93(3)1-28.

Gibbons, R., and K. J. Murphy. 1992. Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence. *Journal of Political Economy* 100(3): 468-505.

Griffith, R., and A. Neely. 2009. Performance Pay and Managerial Experience in Multitask Teams: Evidence from Within a Firm. *Journal of Labor Economics* 27(1): 49-82.

Hazucha, J. F., S. A. Hezlett, and R. J. Schneider. 1993. The Impact of 360-degree Feedback on Management Skills Development. *Human Resource Management* 32(2-3): 325.

Holmstrom, B., and P. Milgrom. 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law Economics & Organization* 7(Special Issue): 24-52.

Kachelmeier, Thornock, and Williamson. 2016. Communicated Values as Informal Controls: Promoting Quality While Undermining Productivity. *Contemporary Accounting Research* 33(4): 1411-1434.
Kelly, N. 2019. Lessons Learned from Measuring International Customer Satisfaction. HubSpot Research. Online article published on February 18, 2019: https://blog.hubspot.com/news-trends/international-nps-research.
Kim, B., Y. K. Chang, and T. Kim. (2018). How Does Corporate Social Responsibility Promote Innovation? The Sequential Mediating Mechanism of Employees' Meaningfulness of Work and Intrinsic Motivation. Proceedings of the 51st Hawaii International Conference on System Sciences.
Kuvaas, B. (2006). Performance appraisal satisfaction and employee outcomes: mediating and moderating roles of work motivation. International Journal of Human Resource Management, 17(3), 504–522. https://doi.org/10.1080/09585190500521581
Lencioni. (2002). Make Your Values Mean Something. *Harvard Business Review*, 80(7), 113–126.
Liu, N., S. Liu, and H. Chu. 2019. Estimating the Effect of a Fit-focused Employee Selection Program. *Journal of Management Accounting Research*, 31(2): 159-175.
London, M., and R. W. Beatty. 1993. 360-Degree Feedback as a Competitive Advantage. *Human Resource Management* 32(2-3): 353-372.
London, M., and J. Smither. 1995. Can Multi-Source Feedback Change Perceptions of Goal Accomplishment Self-Evaluations, and Performance-Related Outcomes? Theory-Based Applications and Directions for Research. *Personnel Psychology* 48(4): 803-839.
London, M., J. W. Smither, and D. J. Adsit. 1997. Accountability. The Achilles' Heel of MultiSource Feedback. *Group and Organization Management* 22(2): 162-184.
Martin, S. R. (2016). Stories about values and valuable stories: a field experiment of the power of narratives to shape newcomers' actions. *Academy of Management Journal*, 59(5), 1707–1724. https://doi.org/10.5465/amj.2014.0061
Meglino, B.M. 1976. A Theoretical Synthesis of Job Performance and the Evaluative Dimension of Organizational Climate: A Social Psychological Perspective. *The Academy of Management Review* 1(3): 58-65.
O'Reilly, C. A., Chatman, J., & Caldwell, D. F. (1991). People and Organizational Culture: A Profile Comparison Approach to Assessing Person-Organization Fit. *Academy of Management Journal*, 34(3), 487–516. https://doi.org/10.5465/256404
Ouchi, W. G. 1979. A conceptual framework for the design of organizational control mechanisms. *Management Science* 25 (9): 833–848.
Prince, J. B., and E. E. Lawler. 1986. Does Salary Discussion Hurt the Developmental Performance Appraisal? *Organizational Behavior and Human Decision Processes* 37: 357-375.
Reilly, R., J. W. Smither, and N. L.Vasilopoulos. 1996. A Longitudinal Study Of Upward Feedback. *Personnel Psychology* 49(3): 599-612.
Robinson, S. L. (1996). Trust and Breach of the Psychological Contract. *Administrative Science Quarterly*, 41(4), 574–599. https://doi.org/10.2307/2393868
Rosso, B. D., Dekas, K. H., & Wrzesniewski, A. (2010). On the meaning of work: A theoretical integration and review. *Research in Organizational Behavior*, 30, 91–127. https://doi.org/10.1016/j.riob.2010.09.001
Simons, R. 2010. *Seven strategy questions: A simple approach for better execution*. Boston, MA: Harvard Business Review Press.
Simons, R., and N. Kindred. 2012. Henkel: Building a Winning Culture. Harvard Business School Case No. 9-112-060.
Silva, P. 2021. Core Values: What They Are, Why They Matter, and How They Can Transform Your Business. *Forbes*, October 12, 2021. Accessed on 7/6/2022.
Smither, J. W., M. London, N. L. Vasilopoulos, R. R. Reilly, R. E. Millsap, and N. Salvemini. 1995. An Examination of the Effects of an Upward Feedback Program Over Time. *Personnel Psychology* 48(1): 1-34.

Thompson, J., & Bunderson, J. (2003). Violations of principle: Ideological currency in the psychological contract. *The Academy of Management Review*, 28(4), 571–586. https://doi.org/10.5465/amr.2003.10899381

Turnley, W.H. and Feldman, D.C., 2000. The impact of psychological contract violations on exit, voice, loyalty, and neglect. *Human relations*, *52*(7), pp.895-922.

Van den Steen, E. 2010a. On the Origin of Shared Beliefs (and Corporate Culture). *RAND Journal of Economics* 41(4): 617-648.

Van den Steen, E. 2010b. Culture Clash: The Costs and Benefits of Homogeneity. *Management Science* 56(10): 1718-1738.

Walker, A. G., and J. W. Smither. 1999. A Five-Year Study of Upward Feedback: What Managers Do with Their Results Matters. *Personnel Psychology* 52(2): 393-423.

WorldatWork in collaboration with J. Bol, I. Grabner, R. Grasser, S. Loftus, T. Sandino, and K. Sedatole. 2019. Contemporary Performance Evaluation Practices.

Table 1, Panel A: Descriptive Statistics

Variable	N	Mean	S.D.	P25	P50	P75	Min	Max
Ln(Sales)	918	11.968	1.398	11.113	12.053	12.969	6.745	14.867
NPS	372	0.340	0.384	0.080	0.333	0.571	-1.000	1.000
Politeness	371	3.822	0.559	3.500	3.778	4.000	1.000	5.000
Helpfulness	371	3.433	0.600	3.000	3.250	3.500	1.000	5.000
% Invoices with Bundles	918	0.263	0.133	0.164	0.254	0.353	0.000	0.714
Failed Audit	105	0.695	0.463	0.000	1.000	1.000	0.000	1.000
Store Manager Change	918	0.080	0.271	0.000	0.000	0.000	0.000	1.000
Promotion Opportunities	889	0.194	0.120	0.098	0.167	0.235	0.091	0.500
Experienced Manager	839	0.758	0.429	1.000	1.000	1.000	0.000	1.000
Late Payment	918	0.126	0.332	0.000	0.000	0.000	0.000	1.000
Average DSI	918	25.853	9.306	17.725	23.631	32.766	14.957	50.062
Sales Days	918	6.955	0.464	7.000	7.000	7.000	1.000	8.000

Notes: Our sample includes store-week level observations for all variables. Variables *Promotion Opportunities*, *Experienced Manager*, *Average DSI*, and *Late Payment* are defined at the store level and are time invariant. Variables *Promotion Opportunities*, *Late Payment,* and *Average DSI* are calculated over the last three months in the pre-period to ensure sufficient recency with respect to the date of the intervention. Observations for the net promoter score (*NPS*) were not available for the entire pre-period, as these metrics were introduced only three weeks before the intervention. The low N for the variable Failed Audit arises from the fact that random audits are performed every week, but not all stores are audited every week. All variables are defined in Appendix 4.

Table 1, Panel B: Covariate Balance

	Control Stores		Treatment Stores		Difference
Variable	**N**	**Mean**	**N**	**Mean**	
Ln(Sales)	288	11.639	288	12.320	0.681***
NPS	40	0.128	43	0.175	0.047
Politeness	40	4.330	42	4.417	0.087
Helpfulness	40	4.258	42	4.202	-0.056
% Invoices with Bundles	288	0.283	288	0.275	-0.008
Failed Audit	30	0.933	31	0.774	-0.159*
Sales Days	288	6.965	288	6.979	0.014
Promotion Opportunities^	15	0.195	16	0.195	0.000
Experienced Manager^	13	0.692	16	0.813	0.120
Late Payment^	16	0.125	16	0.125	0.000
Average DSI^	16	28.200	16	23.763	-4.437

Notes: Table 1, Panel B reports tests of covariate balance between control stores and treatment stores in the pre-period. Two-tailed statistical significance is indicated, respectively, with: * = (p<0.10); ** = (p<0.05); *** = (p<0.01). All variables are defined in Appendix 4. The number of observations varies with the availability of data. Customer surveys began only three weeks before the intervention (and two stores did not have customer surveys in the pre-period), and each store is audited only infrequently (not all stores were audited in the pre-period). One store did not have attendance data available so we could not construct *Promotion Opportunities*, and three control stores did not have an official store manager.

^ We report only one observation per store since these variables are time-invariant. The variables *Promotion Opportunities*, *Late Payment,* and *Average DSI* are calculated over the last three months in the pre-period to ensure sufficient recency with respect to the date of the intervention.

Table 2: Correlation Table

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	*Ln(Sales)*	1.000						
(2)	*NPS*	-0.043	1.000					
(3)	*Politeness*	-0.009	0.105**	1.000				
(4)	*Helpfulness*	-0.038	-0.028	0.807***	1.000			
(5)	*% Invoices with Bundles*	0.423***	-0.051	0.001	0.040	1.000		
(6)	*Failed Audit*	0.206**	-0.105	0.295*	0.323**	0.115	1.000	
(7)	*Store Manager Change*	-0.174***	0.043	-0.111**	-0.151***	-0.219***	-0.389***	1.000
(8)	*Promotion Opportunities*	-0.563***	0.130**	0.085	0.055	-0.379***	-0.107	0.131***
(9)	*Experienced Manager*	0.043	0.026	0.014	0.009	0.114***	0.020	0.117***
(10)	*Late Payment*	-0.026	0.006	0.026	0.033	-0.188***	0.016	0.034
(11)	*Average DSI*	-0.564***	-0.061	0.059	0.061	-0.283***	0.030	0.129***
(12)	*Sales Days*	0.202***	0.020	-0.295***	-0.254***	0.138***	-0.013	-0.076**

		(8)	(9)	(10)	(11)
(8)	*Promotion Opportunities*	1.000			
(9)	*Experienced Manager*	-0.120***	1.000		
(10)	*Late Payment*	-0.251***	-0.008	1.000	
(11)	*Average DSI*	0.187***	-0.261***	-0.060*	1.000
(12)	*Sales Days*	-0.186***	-0.014	0.030	-0.022

Notes: Table 2 reports the pairwise Pearson correlation coefficients across all variables of interest in this study. Two-tailed statistical significance is indicated, respectively, with: * = ($p<0.10$); ** = ($p<0.05$); *** = ($p<0.01$). All variables are defined in Appendix 4.

Table 3: Effects of the Introduction of the Values-Based PA System on Alignment with Core Values
This table reports results from running the following regression:

$$Value\ Alignment_{i,t} = \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Store\ Manager\ Change_{i,t} + \beta_3 Sales\ Days_{i,t} + \beta_n(Store\ Fixed\ Effects) + \gamma_t(Week\ Fixed\ Effects) + \varepsilon$$

Core Values	*We Gain Control of Our Career by Working Hard*	*We Give More Value*		*We are Honest and Ethical*	*We are Caring and Respectful*	
Outcomes	*Ln(Sales)*	*NPS*	*% Invoices with Bundles*	*Failed Audit*	*Politeness*	*Helpfulness*
Post*Treatment	**0.201**	**0.013**	**0.002**	**-0.071**	**-0.095**	**0.057**
	(1.55)	**(0.11)**	**(0.09)**	**(-0.36)**	**(-0.58)**	**(0.32)**
Store Manager Change	-0.108	-0.076	-0.025	-0.189	-0.101	-0.103
	(-1.25)	(-0.96)	(-1.19)	(-0.74)	(-1.55)	(-1.52)
Sales Days	0.415***	-0.036	0.018*	0.736**	0.096	0.259**
	(3.50)	(-0.25)	(1.82)	(2.61)	(0.66)	(2.72)
Intercept	8.814***	0.496	0.115	-4.442**	3.325***	2.018**
	(11.08)	(0.49)	(1.47)	(-2.19)	(3.08)	(2.64)
Store FE?	YES	YES	YES	YES	YES	YES
Week FE?	YES	YES	YES	YES	YES	YES
N	918	372	918	98	371	371
# of stores	32	30	32	26	30	30
adj. R^2	0.303	0.067	0.132	0.411	0.344	0.578
adj. R^2 (alt. est.)	0.935	0.092	0.648	0.371	0.347	0.566

Notes: We use a difference-in-differences specification and estimate regression coefficients using OLS with standard errors clustered by store. In all cases, t-statistics are reported in parentheses underneath the corresponding estimated coefficient. Two-tailed statistical significance is indicated, respectively, with: * = (p<0.10); ** = (p<0.05); *** = (p<0.01). Our estimations are performed using the Stata procedure *xtreg*, with fixed effects and standard errors clustered at the store level. While this procedure yields appropriate standard errors, the R^2 is generally underestimated. In the last row we report the R^2 relative to the estimation of the same model using the Stata procedure *areg*, which fits a linear regression absorbing the categorical factor *Store*, which yields a more realistic R^2. We estimate the regression in which *Failed Audit* is the dependent variable using a linear probability model. Because the assignment of a store to the treatment versus control group is time-invariant, the inclusion of store fixed effects causes the coefficient relative to the variable *Treatment* not to be estimated, hence we are not reporting a row for this variable. We do not report a main effect of *Post* as it is absorbed by the week fixed effects. All variables are defined in Appendix 4.

Table 4: Influence of Promotion Opportunities on the Value-Alignment Effects of the Introduction of the Values-Based PA System

This table reports results from running the following regression:

$$Value\ Alignment_{i,t} = \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Post_t * Promotion\ Opportunities_i + \beta_3 Post_t * Treatment_i * Promotion\ Opportunities_i + \beta_4 Store\ Manager\ Change_{i,t} + \beta_5 Sales\ Days_{i,t} + \beta_n(Store\ Fixed\ Effects) + \gamma_t(Week\ Fixed\ Effects) + \varepsilon$$

Core Values	*We Gain Control of Our Career by Working Hard*	*We Give More Value*		*We are Honest and Ethical*	*We are Caring and Respectful*	
Outcomes	*Ln(Sales)*	*NPS*	*% Invoices with Bundles*	*Failed Audit*	*Politeness*	*Helpfulness*
*Post*Treatment*	0.259**	0.077	0.006	-0.058	-0.058	0.082
	(2.72)	(0.69)	(0.21)	(-0.29)	(-0.34)	(0.43)
*Post*Promotion Opportunities*	-2.116***	-0.666	-0.104	-0.560	0.145	0.337
	(-6.72)	(-1.22)	(-0.91)	(-0.63)	(0.23)	(0.42)
Post*Treatment* Promotion Opportunities	**3.527*** (6.13)**	**1.670** (2.29)**	**0.435* (1.74)**	**-0.102 (-0.04)**	**-0.029 (-0.02)**	**-0.220 (-0.15)**
Store Manager Change	-0.153*	-0.091	-0.043*	-0.157	-0.092	-0.096
	(-1.85)	(-1.04)	(-1.85)	(-0.40)	(-1.44)	(-1.57)
Sales Days	0.353**	-0.079	0.006	0.693**	0.098	0.263***
	(2.68)	(-0.54)	(1.14)	(2.42)	(0.66)	(2.80)
Intercept	9.282***	0.826	0.209***	-4.156*	3.354***	2.017**
	(10.16)	(0.81)	(5.21)	(-2.03)	(3.07)	(2.63)
Store FE?	YES	YES	YES	YES	YES	YES
Week FE?	YES	YES	YES	YES	YES	YES
N	889	360	889	98	359	359
# of stores	31	29	31	26	29	29
adj. R^2	0.360	0.075	0.152	0.400	0.349	0.578
adj. R^2 (alt. est.)	0.943	0.077	0.672	0.349	0.352	0.565

Notes: We use a difference-in-differences specification and estimate regression coefficients using OLS with standard errors clustered by store. In all cases, t-statistics are reported in parentheses underneath the corresponding estimated coefficient. Two-tailed statistical significance is indicated, respectively, with: * = (p<0.10); ** = (p<0.05); *** = (p<0.01). Our estimations are performed using the Stata procedure *xtreg*, with fixed effects and standard errors clustered at the store level. While this procedure

yields appropriate standard errors, the R^2 is generally underestimated. In the last row we report the R^2 relative to the estimation of the same model using the Stata procedure *areg*, which fits a linear regression absorbing the categorical factor *Store*, which yields a more realistic R^2. The variable *Promotion Opportunities* is calculated as (# stores within a 1-mile radius) / (# staff members at the store plus staff members at the stores in a one-mile radius), where the number of store staff members is captured at the median level in the last three months in the pre-period. For ease of interpretation, we have mean-centered the variable *Promotion Opportunities*. Because the assignment of a store to the treatment versus control group is time-invariant, the inclusion of store fixed effects causes the coefficient relative to the variable *Treatment* and *Promotion Opportunities* not to be estimated, hence we are not reporting a row for these variables and their interaction. We do not report a main effect of *Post* as it is absorbed by the week fixed effects. All variables not defined here, are defined in Appendix 4.

Table 5: Influence of Store Manager Experience (Exposure to Values) on the Value-Alignment Effects of the Introduction of the Values-Based PA Assessment System

This table reports results from running the following regression:

$$Value\ Alignment_{i,t} = \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Post_t * Experienced\ Manager_i + \beta_3 Post_t * Treatment_i * Experienced\ Manager_i + \beta_4 Store\ Manager\ Change_{i,t} + \beta_5 Sales\ Days_{i,t} + \beta_n (Store\ Fixed\ Effects) + \gamma_t (Week\ Fixed\ Effects) + \varepsilon$$

Core Values	*We Gain Control of Our Career by Working Hard*	*We Give More Value*		*We are Honest and Ethical*	*We are Caring and Respectful*	
Outcomes	*Ln(Sales)*	*NPS*	*% Invoices with Bundles*	*Failed Audit*	*Politeness*	*Helpfulness*
*Post*Treatment*	-0.253	-0.000	0.075	0.020	-0.747***	-0.839***
	(-1.31)	(-0.00)	(1.00)	(0.05)	(-3.14)	(-3.11)
*Post*Experienced Manager*	-0.437**	-0.053	0.033	0.204	-0.488*	-0.612**
	(-2.43)	(-0.30)	(0.65)	(0.95)	(-1.86)	(-2.06)
Post*Treatment* Experienced Manager	**0.515****	**-0.031**	**-0.109**	**-0.157**	**0.842****	**1.215*****
	(2.18)	**(-0.15)**	**(-1.37)**	**(-0.32)**	**(2.75)**	**(3.70)**
Store Manager Change	-0.073	-0.079	-0.029	-0.249	-0.067	-0.068
	(-0.89)	(-0.88)	(-1.47)	(-0.72)	(-1.10)	(-1.32)
Sales Days	0.402***	-0.117	0.019*	0.719**	0.050	0.267*
	(3.23)	(-0.90)	(1.83)	(2.08)	(0.31)	(2.05)
Intercept	9.092***	1.066	0.120	-4.304	3.651***	1.959*
	(10.94)	(1.20)	(1.49)	(-1.72)	(3.08)	(2.02)
Store FE?	YES	YES	YES	YES	YES	YES
Week FE?	YES	YES	YES	YES	YES	YES
N	839	350	839	88	349	349
# of stores	29	27	29	23	27	27
adj. R^2	0.354	0.090	0.133	0.447	0.355	0.599
adj. R^2 (alt. est.)	0.953	0.122	0.664	0.393	0.359	0.590

Notes: We use a difference-in-differences specification and estimate regression coefficients using OLS with standard errors clustered by store. Two-tailed statistical significance is indicated, respectively, with: * = (p<0.10); ** = (p<0.05); *** = (p<0.01). Our estimations are performed using the Stata procedure *xtreg*, with fixed effects and standard errors clustered at the store level. While this procedure yields appropriate standard errors, the R^2 is generally underestimated. In the last row

we report the R^2 relative to the estimation of the same model using the Stata procedure *areg*, which fits a linear regression absorbing the categorical factor *Store*, which yields a more realistic R^2. The variable *Experienced Manager* is an indicator variable equal to one if the store manager had been in his role at least six months at the time of the intervention, and zero otherwise. Because the assignment of a store to the treatment versus control group is time-invariant, the inclusion of store fixed effects causes the coefficient relative to the variables *Treatment* and *Experienced Manager* not to be estimated, hence we are not reporting a row for these variables and their interaction. We do not report a main effect of *Post* as it is absorbed by the week fixed effects. All variables not defined here, are defined in Appendix 4.

Table 6: Influence of Late Payments on the Value-Alignment Effects of the Introduction of the Values-Based PA System

This table reports results from running the following regression:

$$Value\ Alignment_{i,t} = \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Post_t * Late\ Payment_i + \beta_3 Post_t * Treatment_i * Late\ Payment_i + \beta_4 Store\ Manager\ Change_{i,t} + \beta_5 Sales\ Days_{i,t} + \beta_n (Store\ Fixed\ Effects) + \gamma_t (Week\ Fixed\ Effects) + \varepsilon$$

Core Values	*We Gain Control of Our Career by Working Hard*	*We Give More Value*		*We are Honest and Ethical*	*We are Caring and Respectful*	
Outcomes	*Ln(Sales)*	*NPS*	*% Invoices with Bundles*	*Failed Audit*	*Politeness*	*Helpfulness*
*Post*Treatment*	0.255*	0.083	0.013	-0.055	-0.055	0.042
	(1.83)	(0.67)	(0.44)	(-0.21)	(-0.31)	(0.21)
*Post*Late Payment*	0.049	-0.038	0.073***	0.008	-0.163	-0.366
	(0.25)	(-0.38)	(3.50)	(0.03)	(-1.09)	(-1.65)
Post*Treatment*	**-0.438**	**-0.532*** **	**-0.080**	**-0.164**	**-0.312**	**0.069**
Late Payment	**(-1.60)**	**(-3.85)**	**(-1.30)**	**(-0.38)**	**(-1.34)**	**(0.23)**
Store Manager Change	-0.124	-0.083	-0.031	-0.204	-0.098	-0.086
	(-1.47)	(-1.01)	(-1.65)	(-0.74)	(-1.46)	(-1.32)
Sales Days	0.417***	-0.049	0.018	0.766**	0.088	0.257***
	(3.57)	(-0.35)	(1.70)	(2.33)	(0.61)	(2.80)
Intercept	8.801***	0.584	0.120	-4.653*	3.381***	2.024**
	(11.24)	(0.59)	(1.48)	(-1.98)	(3.17)	(2.76)
Store FE?	YES	YES	YES	YES	YES	YES
Week FE?	YES	YES	YES	YES	YES	YES
N	918	372	918	98	371	371
# of stores	32	30	32	26	30	30
adj. R^2	0.313	0.088	0.140	0.397	0.350	0.583
adj. R^2 (alt. est.)	0.936	0.113	0.652	0.345	0.352	0.571

Notes: We use a difference-in-differences specification and estimate regression coefficients using OLS with standard errors clustered by store. In all cases, t-statistics are reported in parentheses underneath the corresponding estimated coefficient. Two-tailed statistical significance is indicated, respectively, with: * =

(p<0.10); ** = (p<0.05); *** = (p<0.01). Our estimations are performed using the Stata procedure *xtreg*, with fixed effects and standard errors clustered at the store level. While this procedure yields appropriate standard errors, the R^2 is generally underestimated. In the last row we report the R^2 relative to the estimation of the same model using the Stata procedure *areg*, which fits a linear regression absorbing the categorical factor *Store*, which yields a more realistic R^2. The variable *Late Payment* is an indicator variable equal to 1 if the store manager of the store was paid after the half of the close stores (i.e., stores in a one-mile radius) at least one time in the last three months in the pre-period, and zero otherwise. Because the assignment of a store to the treatment versus control group is time-invariant, the inclusion of store fixed effects causes the coefficient relative to the variables *Treatment* and *Late Payment* not to be estimated, hence we are not reporting a row for these variables and their interaction. We do not report a main effect of *Post* as it is absorbed by the week fixed effects. All variables not defined here, are defined in Appendix 4.

Table 7: Influence of Days Sales in Inventory (DSI) on the Value-Alignment Effects of the Introduction of the Values-Based PA System

This table reports results from running the following regression:

$$Value\ Alignment_{i,t} = \alpha + \beta_1 Post_t * Treatment_i + \beta_2 Post_t * Low\ DSI_i + \beta_3 Post_t * Treatment_i * Low\ DSI_i + \beta_4 Store\ Manager\ Change_{i,t} + \beta_5 Sales\ Days_{i,t} + \beta_n(Store\ Fixed\ Effects) + \gamma_t(Week\ Fixed\ Effects) + \varepsilon$$

Core Values	*We Gain Control of Our Career by Working Hard*	*We Give More Value*		*We are Honest and Ethical*	*We are Caring and Respectful*	
Outcomes	*Ln(Sales)*	*NPS*	*% Invoices with Bundles*	*Failed Audit*	*Politeness*	*Helpfulness*
*Post*Treatment*	0.270*	0.054	0.042	0.121	-0.092	0.074
	(1.70)	(0.46)	(1.41)	(0.52)	(-0.46)	(0.32)
*Post*Low DSI*	0.186	-0.080	0.077***	0.301	-0.083	0.019
	(1.01)	(-0.29)	(3.42)	(1.00)	(-0.36)	(0.09)
Post*Treatment*	**-0.317**	**-0.143**	**-0.174*** **	**-0.831**	**0.003**	**-0.065**
Low DSI	**(-1.49)**	**(-0.45)**	**(-4.59)**	**(-1.55)**	**(0.01)**	**(-0.21)**
Store Manager Change	-0.104	-0.085	-0.027	-0.246	-0.106	-0.103
	(-1.18)	(-1.05)	(-1.22)	(-0.91)	(-1.63)	(-1.45)
Sales Days	0.414***	-0.050	0.018*	0.720**	0.090	0.257***
	(3.53)	(-0.35)	(1.72)	(2.56)	(0.62)	(2.81)
Intercept	8.820***	0.596	0.117	-4.319**	3.366***	2.034***
	(11.21)	(0.59)	(1.44)	(2.14)	(3.13)	(2.77)
Store FE?	YES	YES	YES	YES	YES	YES
Week FE?	YES	YES	YES	YES	YES	YES
N	918	372	918	98	371	371
# of stores	32	30	32	26	30	30
adj. R^2	0.307	0.070	0.178	0.429	0.341	0.576
adj. R^2 (alt. est.)	0.935	0.094	0.677	0.380	0.343	0.564

Notes: We use a difference-in-differences specification and estimate regression coefficients using OLS with standard errors clustered by store. In all cases, t-statistics are reported in parentheses underneath the corresponding estimated coefficient. Two-tailed statistical significance is indicated, respectively, with: * =

(p<0.10); ** = (p<0.05); *** = (p<0.01). Our estimations are performed using the Stata procedure *xtreg*, with fixed effects and standard errors clustered at the store level. While this procedure yields appropriate standard errors, the R2 is generally underestimated. In the last row we report the R^2 relative to the estimation of the same model using the Stata procedure *areg*, which fits a linear regression absorbing the categorical factor Store, which yields a more realistic R^2. The variable *Low DSI* is an indicator assuming value 1 if the individual store fell in the lowest quintile of days sales in inventory in the last three months in the pre-period, and 0 otherwise. Because the assignment of a store to the treatment versus control group is time-invariant, the inclusion of store fixed effects causes the coefficients relative to the variables *Treatment* and *Low DSI* not to be estimated, hence we are not reporting a row for these variables and their interaction. We do not report a main effect of *Post* as it is absorbed by the week fixed effects. All variables not defined here, are defined in Appendix 4.