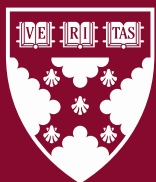


Working Paper 18-069

# Using a 360-Degree Assessment System to Promote Core Values: A Field Experiment in a Retail Chain

Carolyn Deller  
Susanna Gallani  
Tatiana Sandino



**Harvard  
Business  
School**

# Using a 360-Degree Assessment System to Promote Core Values: A Field Experiment in a Retail Chain

Carolyn Deller  
University of Pennsylvania

Susanna Gallani  
Harvard Business School

Tatiana Sandino  
Harvard Business School

**Working Paper 18-069**

Copyright © 2018, 2019, 2020, 2021, 2022 by Carolyn Deller, Susanna Gallani, and Tatiana Sandino.

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Funding for this research was provided in part by Harvard Business School.

## Using a 360-Degree Assessment System to Promote Core Values: A Field Experiment in a Retail Chain

*Carolyn Deller*  
The Wharton School  
University of Pennsylvania

*Susanna Gallani*  
Harvard Business School

*Tatiana Sandino*  
Harvard Business School

January 2022

### **Abstract:**

We conducted a field experiment at a retailer to analyze the effects of introducing a values-based 360-degree assessment system in a setting with pre-existing high-powered monetary incentives linked to financial performance. The retail chain's managing director intended to use this system to encourage her managers not only to pursue values related to financial short-term goals but also values capturing the organization's long-term goals. We explore the effect of the intervention and find that the values-based 360-degree system had more favorable effects for stores with greater promotion opportunities and for stores with managers with experience related to the implementation of values. We also find that management behavior that could be perceived to be inconsistent with the spirit of the system (i.e., delaying salary payments) had a detrimental effect on the intervention, leading to worse performance. Our findings highlight important factors for successfully implementing values-based 360-degree systems as complements to explicit incentives.

**Keywords:** 360-degree assessments, values, implicit incentives, field experiment, 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 express consent of the organization's legal representatives.

We thank Markus Arnold, Pablo Casas-Arce, Tony Davila, John Harry Evans, Mario Schabus, and participants at the 2021 University of Central Florida Workshop Series, 2019 Information, Markets, and Organizations Conference at Harvard Business School, 2019 AIMA World Conference on Management Accounting Research, 2019 Consortium for Operational Excellence in Retailing, 2019 Cherry Blossom Conference at George Washington University, 2018 AAA Annual Meeting, 2018 Emerging Management Accounting Scholars Symposium, 2018 Global Management Accounting Research Symposium, 2018 Management Accounting Section Midyear Meeting, Erasmus University, and the Harvard Business School 2017 Brown Bag Series for their useful comments and suggestions. We thank Kyle Thomas for his excellent assistance at the research site, as well as in the data collection and analyses, Mayuresh Kumar and Gunjan Kapil for their research assistance at the research site, Trang Nguyen for research assistance in the later stages of 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 field experiment and providing us with access to data. All errors are our own.

## 1. INTRODUCTION

In this paper, we study the implementation of a *values-based 360-degree performance assessment system*, that is, one designed to communicate and reinforce an organization's core values and to provide a complete (360-degree) evaluation from multiple raters: the ratee, his or her subordinates, supervisors, peers, customers, and so on (London and Smither 1995). We do so in a setting with pre-existing high-powered incentives explicitly linked to financial results (hereafter, high-powered explicit incentives). High-powered explicit incentives can motivate employees to behave entrepreneurially but can also lead to an excessive focus on short-term, individual-wealth-maximizing activities at the expense of values associated with long-term organizational goals (Holmstrom and Milgrom 1991; Baker, Gibbons, and Murphy 1994). Many organizations use performance assessment systems promoting not only short-term results but also long-term organizational values. In a WorldatWork survey of 254 human resource managers from a diverse set of firms, 72% of the respondents indicated that their company's performance evaluation process reflected organizational core values to a significant extent.<sup>1</sup> This percentage was even higher (80%) among respondents from companies relying on performance evaluations incorporating multiple raters (48% of the survey respondents). We study whether and under what conditions a 360-degree system promoting corporate values led to changes in performance and behaviors consistent with those values and actionable within a short timeframe.

Research has examined the performance effects of 360-degree systems by comparing initial ratings to subsequent ratings and found only modest improvements (e.g., Smither, London, and Reilly 2005). Our study extends this line of research by: (a) examining the introduction of a *values-based 360-degree system* in a common but previously unexamined setting—that is, in an organization with high-powered explicit

---

<sup>1</sup> This is the percentage of respondents indicating that the extent to which their company's performance evaluation process reflected 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 included organizations from multiple industries and different sizes.

incentive contracts implementing a 360-degree system not linked explicitly to rewards<sup>2</sup>; (b) uncovering and testing conditions under which 360-degree systems are more or less likely to be effective (a gap in the literature identified by scholars such as Smither et al. 2005); and (c) examining the effects of a 360-degree system on independent performance measures<sup>3</sup> linked to the behaviors emphasized by the 360-degree system.

A values-based 360-degree performance assessment system not linked to monetary rewards could affect employee behaviors through two mechanisms identified in the performance assessments literature (e.g., Boswell and Boudreau 2002): (a) an *evaluative* mechanism providing implicit incentives, especially to employees interested in pursuing a career within the organization; and (b) a *developmental* mechanism through which employees (both raters and ratees) learn (more) about desired behaviors that create value in the organization. Conversely, in the presence of high-powered explicit incentives, values that are included in a 360-degree system but not linked directly to pay may be ignored. We examine the overall effectiveness of such a system in shaping employee behavior. More importantly, we examine conditions that may strengthen the underlying mechanisms, thereby affecting the outcomes of the implementation.

For this study, we partnered with a growing retail chain in a major city in India that had traditionally compensated store employees with high-powered explicit incentives based on financial performance. While congruent with some of the firm's overall values—specifically those related to “working hard”—these incentives had also introduced behaviors (such as deceiving customers and gaming the system) that were detrimental to other firm values aligned with the organization's long-term goals. When the company was sufficiently small, the managing director could personally shape and monitor the behaviors of employees through direct informal communication of her vision and values and by working alongside store teams in

---

<sup>2</sup> In fact, scholars have emphasized the role of 360-degree systems in shaping organizational culture and introducing/reinforcing corporate values (London and Beatty 1993; London and Smither 1995) but, to the best of our knowledge, have yet to empirically examine this important role.

<sup>3</sup> By independent performance measures we refer to metrics that are captured by information systems other than the 360-degree employee survey. These metrics are partial indicators of whether the 360-degree assessment system was effective on a subset of observable behaviors among the broader set of behaviors promoted by the values-based 360-degree system.

their stores. However, as the company scaled up, she needed a more formal system to communicate and uphold the company's vision and values.<sup>4</sup> Management therefore introduced a values-based 360-degree assessment system in tandem with a formalized vision and core values.

The 360-degree system was not tied to any new monetary incentives and was centered on the store manager. Appraisals were based on surveys evaluating store managers and were completed by the store managers themselves, their store team members, and their supervisors. The survey questions were arranged by the company's core values, capturing the behaviors desired by senior management. Appraisals also incorporated information from a separate survey used to capture customers' experiences at the store. The goal was to emphasize employee development by communicating corporate values to raters and ratees, and to highlight desired behaviors to store managers, who were, in turn, expected to help their teams develop desired behaviors and performance. A secondary objective was to communicate what behaviors mattered for career advancement and to identify employees suitable for promotion. Existing monetary incentives, unrelated to the 360-degree system, were kept.

We conducted a natural field experiment where we introduced the 360-degree system (as well as the core values and vision) in a randomly selected treatment group comprising half of the chain's stores in order to test its effectiveness.<sup>5</sup> We use a difference-in-differences approach to compare the performance of treated stores with that of control stores and examine conditions moderating the evaluative and developmental mechanisms expected to drive system effectiveness. These conditions are: (1) heterogeneity in *promotion opportunities* across business units (in this case, stores), potentially affecting the strength of the evaluative mechanism associated with the system; and (2) variation in the *business unit managers' experience* (in this

---

<sup>4</sup> We describe the core values in detail in Section 3.2.

<sup>5</sup> We categorize our study as a natural field experiment because the subjects did not know they were participating in an experimental study. Instead, they were told the 360-degree system was introduced in a first set of stores (as a first phase) and was expected to continue to be used in their stores for the following months. This design provided us with significant advantages as it allowed us to avoid subject self-selection into the experiment, to conduct the study in the natural environment where the subjects worked, to provide the subjects with a sense of continuity, and to draw causal inferences from the intervention, due to the randomized selection of the treatment group (Floyd and List 2016).

case, regarding how to behave according to corporate values), potentially affecting the strength of the developmental mechanism of the system.

Our sample period ends *before* any feedback was provided to store managers at treated stores;<sup>6</sup> thus, rather than documenting the feedback effects of a 360-degree values-based system, we examine its effects on the behavior of employees formally learning about the core values; completing 360-degree surveys incorporating those values; and anticipating being evaluated on those values in the future.<sup>7</sup> Our main dependent variables are company–values–related metrics capturing dimensions of performance that could be influenced by the store staff within a short timeframe (i.e., sales, customer service, honesty, and treatment of others).

We find no significant average effects attributable to the intervention on any of the measures associated with the company’s core values. However, as predicted, we find that the 360 degree-system worked better where the mechanisms expected to influence the effectiveness of the evaluation system were likely to be stronger. First, the intervention had a more positive effect in stores where the *evaluative* aspect of the 360-degree system was likely to be more salient. Specifically, we found that greater opportunities for promotion were associated with significant marginal increases in sales (consistent with the system promoting a core value related to hard work), in the likelihood of customers recommending the retail chain (captured with Net Promoter Score (NPS)), and in the percentage of invoices including bundles (these latter two results are consistent with the system promoting a core value related to giving more value to customers), in treated stores relative to control stores, following the 360-degree intervention. Second, the values-based 360-degree system had a more significant effect on behaviors associated with the core values related to giving more value to customers and being caring and respectful where the *developmental* aspect of the 360-degree system was likely to be more effective. We find that in treated stores with managers that had prior experience with core values (and thus greater ability to prioritize and model value-related behaviors and

---

<sup>6</sup> We provide a justification for this design choice in section 3.3.

<sup>7</sup> While the store managers expected to be evaluated again based on the 360-degree system in the future, other raters in the system might have expected that these values would be important criteria considered for their career progression.

facilitate their development within their units' teams), the intervention was associated with significant marginal increases in the average scores provided by customers when rating their *Satisfaction*, and the *Politeness* and *Helpfulness* of the store staff, compared to control stores, following the intervention.

Our interviews also revealed conditions that could have led to a less effective intervention, and may explain the insignificant average effect of the intervention. Interviews suggested that the store teams may have been less receptive to the intervention where management's behavior did not seem to align with the core values underlying the 360-degree system. While the core values emphasized providing great value to customers, some salespersons claimed that management did not provide them with enough inventory to do so. In addition, store teams were expected to be honest and trustworthy, but some salespersons questioned the company's adherence to these values as they did not always receive their monthly compensation on time. We exploited variation in inventory levels and in late payments to examine the impact of these factors on the effectiveness of the intervention. We find that treated stores ranking in the lowest quintile of pre-intervention average days-of-sales in inventory (DSI), had a marginally lower percentage of invoices including bundles relative to control stores, following the introduction of the 360-degree system. Additionally, treated stores that had been paid late in the pre-intervention period exhibited a significant marginal decline in sales, *NPS*, and in the *Knowledge* and the *Politeness* ratings, relative to control stores, following the implementation of the 360-degree system. These results suggest that a values-based 360-degree system is likely to demotivate rather than motivate employees to incorporate core values in everyday operations if those employees perceive that the company is not keeping its end of the bargain.

We offer several contributions to the literature and to the practice of management accounting. First, as explained earlier, we extend academic knowledge of the effects of 360-degree performance assessment systems. Second, we contribute to the literature on mechanisms to improve performance without explicit monetary incentives. We find conditions where a values-based 360-degree system is more likely to have a positive incremental effect on financial and non-financial performance in the absence of additional tangible rewards. Prior work has examined other mechanisms unrelated to monetary incentives that may motivate employees to improve performance, such as transparent performance feedback (Bernstein and Li 2017),



relative performance information (Blanes i Vidal and Nossol 2011), and recognition (Bradler, Dur, Neckermann, and Non 2016, Gallani 2022).

Third, we contribute to the literature examining the use of multiple performance measures (e.g., Ittner, Larcker, and Randall 2003; Ittner, Larcker, and Meyer 2003; Campbell 2008; Hall 2008), here in the context of a 360-degree assessment system. Our results are consistent with the conjecture that the introduction of a values-based 360-degree system may be more effective under certain conditions, specifically where employees have higher implicit incentives due to promotion opportunities, greater ability and guidance to implement multiple values due to their managers' experience with, and exposure to, core values, and where employees perceive that top managers are also adhering to the espoused values themselves.

The rest of the paper is organized as follows: Section 2 reviews the literature and states our hypotheses. Section 3 describes our field setting and field experiment. Section 4 presents our empirical analyses. Section 5 contains qualitative insights from interviews and supplementary analyses. Section 6 concludes.

## **2. LITERATURE REVIEW AND HYPOTHESES**

The accounting and economics literatures have long suggested that organizations using explicit rewards tied to financial results could improve goal alignment by adding subjective performance assessments (Baker, Jensen and Murphy 1988; Ittner and Larcker 1998; Prendergast 1999). Financial measures have several limitations: they are transaction-oriented, backward-looking, and unable to capture the value of many intangible assets affecting future outcomes (Kaplan and Norton 1992; Merchant and Sandino 2009). Subjective performance assessments can contribute to promoting behaviors – such as teamwork, integrity, and a positive attitude – aligned with the company's mission and values, beyond those promoted by explicit rewards (Baker et al. 1994; Gibbs et al. 2004).

While subjective assessments are often performed solely by an employee's supervisor, many organizations use 360-degree performance assessment systems to gain a more complete view of employee performance by also including self, peer, subordinate, and customer assessments, including closed- and

open-ended questions on important behavioral dimensions reflecting a company's core values (London and Beatty 1993, WorldatWork 2019).

We examine the performance effects of an implementation of a 360-degree system not linked to monetary rewards, designed to state and reinforce organizational values in a setting with preexisting explicit incentives. More importantly, we examine two conditions that could render such a values-based 360-degree system more effective: 1) greater promotion opportunities, and 2) experience with core values.

### **2.1 Effects of a Values-Based 360-degree Assessment System in a Setting with Explicit Incentives**

A well-designed values-based 360-degree system should contribute to aligning employees with corporate values, including those that are difficult to assess using objective measures and/or that are not tied to explicit rewards, through two mechanisms described and defined in the performance assessment literature: an evaluative and a developmental mechanism (e.g., Boswell and Boudreau 2002; Prince and Lawler 1986). The evaluative mechanism of a performance assessment typically involves comparing an employee's performance with expectations or with the performance of other employees, which may implicitly affect the employee's rewards and career prospects (e.g., through discretionary bonuses, promotions, firings, etc.). The evaluative mechanism of a 360-degree system should motivate an employee to exert effort on the dimensions included in the performance assessment. The developmental mechanism of a performance assessment works by communicating the behaviors that employees should adopt to achieve organizational goals, thus encouraging employees to draw on relevant skills they already possess; to acquire, develop, and, where applicable, teach others skills that will improve effectiveness; and to adopt (and for leaders, help their teams adopt) attitudes consistent with the goals and/or values promoted by the system.

The management literature has examined the performance effects of 360-degree systems, typically comparing the ratings received by a manager in the initial survey to subsequent ratings (e.g., Avery 2000; Smither, London, and Reilly 2005) and without considering the existence of explicit incentives used in

tandem with 360-degree systems.<sup>8</sup> Literature reviews have concluded that research has found only modest improvements in ratings over time and not among all those appraised (Smither, London, and Reilly 2005; Atwater, Brett, and Charles 2007). Smither et al. (2005) provide several explanations for these limited improvements, including the crudeness of looking at average ratings.<sup>9</sup> Additionally, subjective ratings are inherently subject to bias, which could hinder their ability to detect actual changes in the performance they are supposed to capture.<sup>10</sup> We overcome these limitations by analyzing the effects of a values-based 360-degree system on changes in independent performance metrics aligned with company values that are actionable in the short-term,<sup>11</sup> rather than on changes in ratings, as in prior research. The extensive use of values-based 360-degree systems in practice suggests that despite limited empirical evidence of their effectiveness, managers expect such systems to be valuable (WorldatWork 2019).

Although a values-based 360-degree assessment system could encourage employees to exert effort on performance dimensions linked to the company's values, employees may not respond to such system in a setting with explicit incentives, where those incentives are not linked to the 360-degree system. Research suggests that incentive pay directs employees' attention towards the tasks explicitly rewarded and away from other tasks (Holmstrom and Milgrom 1991). Employees may be especially unresponsive to a values-based 360-degree system not linked to rewards if their horizon with the company is short and/or their discounting of future expected payoffs is high (Holmstrom and Milgrom 1991; Baker et al. 1988). They may also decide to focus on activities they already know how to perform rather than actions related to newly emphasized values for which they may lack skills (Meglino 1976). Furthermore, pre-existing incentives

---

<sup>8</sup> Prior studies have focused on mid-level managers in traditional hierarchies whose compensation presumably is largely fixed pay.

<sup>9</sup> Smither et al. (2005) raise concerns about the practice of comparing average ratings at two consecutive times since the ratee might have focused on improving performance related to only a few items and the raters might anchor on their initial impressions.

<sup>10</sup> Subordinates typically provide overly lenient ratings of their bosses, fearing retaliation (Smith and Fortunato 2008), and self-ratings tend to be even more lenient and less reliable (e.g., Smither et al. 1995; Atwater et al. 2007). Finding improvements in overly lenient ratings can be hard, not only due to low variation but also because ratings are capped at the maximum allowed by the scale (ceiling effect).

<sup>11</sup> These proxies were captured by information systems other than the 360-degree survey, and are not meant to be comprehensive of all the changes expected from the intervention but are aimed at capturing a few representative behaviors where there should have been a change.

may have been sufficiently effective at inducing significant effort such that employees expect little payoff from increasing effort beyond current levels. Whether a values-based 360-degree system in a setting with explicit incentives will lead to intended behaviors and outcomes is thus an empirical question. However, given that companies implement values-based 360-degree systems in the presence of explicit incentives intending a positive impact, we formalize the following hypothesis:

*Hypothesis 1: The introduction of a 360-degree values-based assessment in a setting with explicit incentives will positively affect behaviors and outcomes consistent with corporate values.*

Values-based 360-degree systems might be more effective under conditions where the mechanisms expected to drive performance are likely to be stronger: (1) where employees have greater promotion opportunities (i.e., when their perceptions of the 360-degree system as an “evaluative mechanism” leading to potential promotions may be accentuated) and (2) where business unit managers possess relevant experience with respect to actioning the core values (i.e., when the “developmental mechanism” is more likely to bear fruit through business unit managers drawing on their existing knowledge to change and develop their own behavior and to provide training and coaching to team members).

## **2.2 Promotion Opportunities Accentuating the Evaluative Effects of a Values-Based 360-degree System**

Subjective assessments not explicitly linked with monetary rewards are likely to give rise to implicit incentives to the extent that employees anticipate some relation between the assessments and managers’ decisions impacting their career, such as salary revisions, job assignments, or promotions (Gibbons and Murphy 1992; Prendergast 1999; Campbell 2008). Career concerns can therefore provide important incentives even in the presence of other explicit incentives, as employees attempt to influence their employers’ or the labor market’s beliefs about their abilities and suitability for higher ranked and/or better paid positions (Gibbons and Murphy 1992). Prior accounting research suggests that individuals who have greater opportunities for advancement are more likely to exert effort on dimensions of performance that are relevant to the company even though they are not linked to individual monetary rewards. For example,

examining a fast-food chain where promotion and demotion decisions were informed by nonfinancial metrics, Campbell (2008) finds that managers in locations with higher opportunities for promotion performed better on those nonfinancial metrics than managers in other locations.

To the extent that a 360-degree system introduces implicit incentives, it could help align employees with corporate values capturing both short-term and long-term organizational objectives, especially if employees seek career development within the organization and the opportunity to be promoted exists. Additionally, the multi-source nature of the 360-degree system further strengthens the implicit incentive contract by providing top management with comprehensive information for career-related decisions (Prendergast 1999; Loughry and Tosi 2008). Based on these arguments, we state the following hypothesis:

*Hypothesis 2: The introduction of a 360-degree values-based assessment in a setting with explicit incentives will be more positively related to behaviors and outcomes consistent with corporate values in business units offering greater opportunities for promotion than in other units.*

Notwithstanding the arguments supporting these predictions, employees motivated by greater opportunities for promotion may have already spontaneously incorporated actions aligned with the company's values prior to the intervention, following behaviors modeled by company leadership.

### **2.3 Manager Experience Accentuating the Developmental Effects of a Values-Based 360-degree System**

One of the main benefits of a values-based 360-degree system is that "it can call attention to important performance dimensions heretofore neglected by an organization, in the process of conveying organizational values" (London and Beatty 1993, p. 361). Such a system communicates what is valued by the company to participating employees (both raters and ratees) and can trigger discussions among employees on how to incorporate corporate values into the diverse aspects of their day-to-day work. The ability of employees to comprehend and implement the core values in their work, however, may vary with the level of experience that they possess in this domain.

Exposure to top management and their decision-making processes is an important source of learning about the company's priorities, core values, and behaviors that are consistent with them (Van den Steen

2010a). For instance, observing the decisions that top management makes when faced with difficult choices—such as whether or not to give up profits in order to address a customer concern that could affect the long-term prospects of the organization—makes clear what values are prioritized (Simons 2010). Thus, business unit managers with relevant experience have greater knowledge on how to implement values. Consequently, they are likely to be best placed to execute core values given greater prominence through the 360-degree system, and to develop not only themselves but also to provide training and coaching to their unit's team members. Prior research suggests that such training and coaching is necessary to effectively enhance employees' understanding of the values conveyed by a multisource performance assessment system (London, Smither, and Adsit 1997). We conjecture that the introduction of a values-based 360-degree system is more likely to effectively enhance employee understanding of the company's values in business units led by managers with greater exposure to and interaction with top management, who know better how to execute on values and make tradeoffs that reflect those values, and who could transfer that expertise to their subordinates.

More generally, prior research has shown that more experienced managers are more likely to lead their units toward making the right tradeoffs to succeed. In a large distribution firm that was introducing a Balanced Scorecard, Griffith and Neely (2009) find that only units with experienced managers could interpret the Balanced Scorecard measures and improve performance. In their words, "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). In the specific context of a 360-degree system, Atwater and Brett (2005) document better results when the workers considered themselves to be more competent.

Based on these arguments, we state our third hypothesis:

*Hypothesis 3: The introduction of a 360-degree values-based assessment in a setting with explicit incentives will be more positively related to behaviors and outcomes consistent with corporate values in business units led by managers who possess relevant experience with respect to how to execute on those corporate values.*

Despite these predictions, there exists the possibility that business units led by managers with relevant experience may have already exhibited behaviors aligned with corporate values, resulting in a muted reaction to the introduction of a system geared towards raising awareness about those values.

### **3. RESEARCH SETTING AND FIELD EXPERIMENT**

Our research site is a mobile phone retail chain in one of India's main cities (hereafter RETAILER). A typical store 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. To emulate the entrepreneurial spirit, sense of ownership, and incentives of local mom-and-pop stores—the chain's main competitors—the 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 sometimes participate in some of its sales incentive plans.

RETAILER seeks to differentiate itself from mom-and-pop stores 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. 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—employees for unethical behavior such as theft and misleading customers.

While a reliance on the managing director's personal interactions with store teams had helped foster a strong company culture, this was not suited to RETAILER's ambitions for expansion. Furthermore, the managing director sought to more strongly encourage store teams, particularly the managers, to focus not only on short-term financial performance (already strongly incentivized by the compensation structure), but also on long-term behaviors needed to build a consistent and profitable brand; for example, building long-term customer relations, providing feedback to team members, and not lowering prices just to make a sale.

The managing director decided to implement a values-based 360-degree assessment system to formally communicate and promote these values in the hopes of shaping desired behaviors to an even greater extent.

### **3.1 Store Managers and the Values-based 360-degree System**

At the time of the study, RETAILER's store managers were incentivized to behave as owners of the stores that they managed. They could earn generous sales commissions (at higher commission 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.<sup>12</sup>

As in most retail chains (Arnold, Palmatier, Grewal and Sharma 2009), RETAILER's store managers played many roles and were considered key drivers of store success: they led and trained their store teams; were responsible for forming good, long-term customer relationships; and communicated inventory and staffing needs to headquarters. They were accountable for the ongoing success of their stores and were expected to model behaviors conducive to the chain's success.

A 360-degree system makes it possible to measure performance with respect to all these roles. 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 than one limited to assessments by the manager's supervisor.<sup>13</sup> As London and Beatty (1993, p. 360) note, "Subordinates ... may have more complete and accurate information about many leadership behaviors than supervisors have."

By implementing a values-based 360-degree assessment system, RETAILER's management sought not only to formalize and communicate the company's core values and their associated behaviors, but also to gain a complete picture (including insights from open-ended questions) of how each store manager was living those values, so as to provide feedback and coaching. The 360-degree surveys were designed to

---

<sup>12</sup> These calculations are based on data shared by the company relating to store manager payments in July of 2013. The structure of the store managers' compensation at that time was similar to that during our sample period.

<sup>13</sup> 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 HR functions, supply chain, etc.) in addition to supervising the stores.



collect information about each store manager from the store's cashier and brand promoters, the manager's supervisor, the customers, and the manager himself.

The system was presented to employees (both raters and ratees) as a tool to learn what the company considered important, as well as to help them grow and increase their chances of promotion (thus, it provided 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 pay) or be promoted to "store manager coach" to assist nearby stores (while retaining their ability to earn high-powered financial rewards).

The hope was that with management (1) formally communicating core values, (2) periodically asking store teams and supervisors to consider those values while assessing their store manager's behavior (or, for the store manager, his own), (3) providing feedback to store managers, and (4) making it clear that those values would be considered for promotion decisions, the teams would internalize both the values that related to achieving short-term financial results, as well as those associated with behaviors that would contribute to building a successful company brand in the long run. This, in turn, would enable more delegation and less direct monitoring by the managing director (for instance, if store teams could be trusted to behave in the interest of the company's long-term success) and greater work satisfaction and motivation for store teams (Van den Steen (2010b)).

### **3.2 Surveys Used as Input to the Values-based 360-degree System**

The surveys associated with the 360-degree system asked each respondent (the store manager, store members, and the supervisor) to assess the manager on various behaviors, organized by the four core values of the firm: (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 also asked where the store manager was doing a good job and where he could do better. The survey instrument is shown in Appendix 1.

One customer per store was randomly selected each day for a telephone survey. They 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 question, a store’s *Net Promoter Score* (hereafter NPS) can be 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).<sup>14</sup> The NPS is a popular metric for customer experience.

### **3.3 The Field Experiment**

Since the managing director was interested in understanding the effectiveness of a values-based 360-degree assessment system (developed and implemented via extensive consultation with the research team), she readily agreed to introduce it as a field experiment, implementing the system in approximately half the stores.<sup>15</sup> While she was enthusiastic, it was unclear whether store employees who had been focused on monetary rewards up to that point would respond to the intervention, or whether they would respond by exerting more effort only on certain performance dimensions (for instance, known behaviors linked to existing incentives or behaviors not linked to incentives but emphasized by the new system).

Store selection was randomized. However, we grouped stores in “blocks” if they were close enough to each other such that contamination effects would be a concern if some 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.<sup>16</sup> Participants in the stores selected for treatment were advised that RETAILER was piloting the system only in certain stores as a first phase of implementing the system throughout the company. During this phase, they were asked not to discuss it outside their store team.

The system was launched with an inspiring presentation and interactive session led by the managing director. The session was held twice, at the end of March 2015 and at the end of April 2015, to include all

---

<sup>14</sup> See [www.netpromoter.com](http://www.netpromoter.com).

<sup>15</sup> 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.

<sup>16</sup> 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.

the treated stores and to help some stores that had difficulties accessing the 360-degree surveys online.<sup>17</sup> The presentation had two parts. The first introduced store teams to the company's competitive strategy and to its newly formalized vision statement and core values (included in Appendix 2), emphasizing the importance of the role of the store managers and store teams to accomplish these values. During the second part, the managing director explained that a 360-degree system was to be implemented at the stores in attendance in order to gain a comprehensive understanding of the support the store manager provided and his commitment to the core values.<sup>18</sup> Attendees were advised that the aggregated survey responses (confidentiality of individual responses was assured) would be used to provide feedback and coaching to store managers. Attendees were then asked to complete the performance appraisal survey (Appendix 1) in a computer lab that was set up for this purpose.<sup>19</sup>

The feedback sessions were held about three months after the launch of the 360-degree system. By this time, a couple of treated store managers had moved to control stores (we speak to this further in footnote 26) and a couple of feedback sessions that should have taken place did not, reducing our available sample for the period post 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). Hence, we end our post-period right before the feedback sessions.<sup>20</sup> This potentially decreases the power of our

---

<sup>17</sup> 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 (6 weeks in total).

<sup>18</sup> As shown in Appendix 1, this survey included 48 items. The number of items on survey instruments for upward feedback (in which subordinates provide feedback to their superior) and 360-degree feedback 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.

<sup>19</sup> Store staff could complete the survey privately. Store team members missing the session were able to complete the survey privately, outside of 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. These favorable ratings suggest 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, and their customers ratings resulted in an average NPS of 33% (which seems acceptable, compared to average NPS scores of 27% in India and 49% in the United States (Kelly 2019)).

<sup>20</sup> The absence of feedback in the post-period that we study may have affected the results that we document, but research suggests this is not a significant issue. Smither, London, Vasilopoulos, Reilly, Millsap and Salvemini (1995) find that managers who received individualized feedback in an upward feedback system were no more likely to improve their performance than managers who did not. (In their setting, managers with fewer than three subordinates received only an aggregate organizational report to protect the subordinates' anonymity). Using the same managers

tests, but it also allows us to examine the motivational anticipatory effects of the 360-degree system implementation, independent of any feedback effects.

## 4. RESEARCH DESIGN AND EMPIRICAL TESTS

### 4.1 Research Design

We test the effects of the intervention using independent<sup>21</sup> performance measures that were readily available (or could be easily constructed) from the organization's existing information system as proxies for the desired behaviors highlighted by the 360-degree assessment survey.<sup>22</sup> The chosen measures of performance are (1) related to the core values emphasized by the 360-degree system, (2) responsive to employee actions in the short-term, (3) more objectively measured than those used in prior studies, and (4) less subject to leniency bias. Furthermore, the availability of a treatment and a control group allows us to better assess causal relations relative to prior studies relying predominantly on initial and subsequent performance assessments under the system itself.

Our empirical analyses, described in the remainder of this section, explore the performance effects of the implementation of a values-based 360-degree system in a setting with explicit incentives.

#### 4.1.1 Data and Measures Analyzed

We examine how treated stores performed vis-à-vis control stores on dimensions capturing behaviors related to the four core values assessed through the 360-degree system.<sup>23</sup> For the first core value, "*We gain control of our own career by working hard every day and reaching out for support,*" we measure

---

as in 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).

<sup>21</sup> Footnote 16 explains the "independence" of these metrics.

<sup>22</sup> All of the measures we examine pertain to information that had historically been recorded with the exception of the customer surveys, which began shortly before the introduction of the 360-degree feedback system and were the result of our collaboration with the company.

<sup>23</sup> In our setting, we cannot disentangle the effects of communicating the company's vision and values from the effect of formalizing them through a 360-degree system, because the intervention included both factors. We encourage future research to explore the relative effectiveness of these two components of the intervention.

performance using  $\ln(\text{Sales})$ , the natural logarithm of store-level weekly sales.<sup>24</sup> For the second core value, “*We give more value*,” we use store-week-level metrics capturing several aspects of the service customers received at the store. These are the net promoter score (*NPS*), customers’ perception of the sales associates’ knowledge (*Knowledge*) and their satisfaction with the interaction with the sales associate (*Satisfaction*), all of which come from the customer surveys, and 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. We measure performance 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 audited 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.<sup>25</sup> Store managers were directly responsible for any unfavorable deviation and were penalized with pay deductions equivalent to the value of the missing assets. 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*). Sales performance was already rewarded with monetary rewards in the pre-existing incentive system, and, in some cases, store teams were financially penalized for significantly 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.

All measures were discussed with the managing director, who approved them as appropriate proxies representative of the underlying dimensions of performance that she intended to emphasize with the 360-degree system. Appendix 2 lists the measures we use to capture these dimensions of performance and their corresponding core values.

To analyze organizational conditions influencing the evaluative and developmental mechanisms underlying the 360-degree system, we define the variable *Promotion Opportunities* as (# stores within a 1-

---

<sup>24</sup> The weekly sales amounts are net of customer returns.

<sup>25</sup> Note that our sample size for our analyses using audit results as the dependent variable is relatively small because not all stores were audited each month, and we dropped any observations pertaining to stores that did not have at least one audit in both the pre- and the post-period. Consequently, we have low power to detect effects.

mile radius) / (# store team members within a 1-mile radius).<sup>26</sup> Additionally, we capture the store manager's experience with the company core values by constructing 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 360-degree system, and zero otherwise. In a retail chain setting, the experience relevant to a values-based 360-degree 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 day-to-day work. In our setting, 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.<sup>27</sup> The managing director considered a six-month period to be long enough for store managers to learn how to behave according to values. All variables used in our empirical tests are defined in Appendix 3.

In late May and early June 2015 (that is, a little over a month after the system launched), we asked an assistant in India (identified by the managing director) to conduct follow-up interviews on our behalf to gauge the sentiment of the employees with respect to the 360-degree system and enhance our interpretation of our quantitative analyses. Fourteen store staff members (two store managers, three cashiers, and nine promoters) from five stores randomly selected from the treated stores were interviewed using a structured interview questionnaire (see Appendix 4).

---

<sup>26</sup> 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 describe the opportunities for promotion in greater detail in section 4.2.2. We defined 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 uses a similar indicator on 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 (2) the number of other employees likely to take that position is low.

<sup>27</sup> To avoid contamination, the managing director did not explicitly discuss the core values or the 360-degree system during the time of the intervention.

Our sample includes weekly observations for 32 stores (16 each in the treatment and control groups) spanning 22 weeks, of which 11 are prior to the intervention (pre-period) and 11 are subsequent to the week when all the store employees affiliated with the treated stores completed the 360-degree surveys (post-period).<sup>28, 29</sup> We 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 694 store-week observations. We have fewer observations for our *NPS*, *Politeness*, *Knowledge*, *Satisfaction*, and *Helpfulness* measures, since the customer surveys began only three weeks before the 360-degree system was first launched, as well as for the *Failed Audit* variable, since stores were randomly audited on average once a month.

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.<sup>30</sup> Measures of customers' perceptions related to the service received (i.e., *Knowledge*, *Helpfulness*, *Politeness*, and *Satisfaction*) span across a 1-5 Likert scale, and present, on average, opportunities for improvement. The *% Invoices with Bundles* is below 34% 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. With few exceptions (only 6.2% of the store-weeks analyzed), stores record sales on every day of the week each week. About 76% of

---

<sup>28</sup> Our initial sample included 39 stores (20 in the treatment group and 19 control stores). Of these, seven were dropped. Two stores that were supposed to be included in the treated group were eliminated 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). Another store was dropped because it opened immediately before the sample period and closed immediately after it. In the case of two other stores, the store team participated only in the core values session and not in the 360-degree survey. One control store was contaminated by the reallocation of a treated store manager immediately after the intervention. One store team identified the cashier as the (acting) store manager, but the cashier himself did not self-identify as store manager in his survey.

<sup>29</sup> Our pre-period includes the 11 weeks preceding the first launch to keep correspondence between the length of the pre-period and the length of the post-period, in line with Charness and Gneezy (2009).

<sup>30</sup> In the analysis of variables obtained from the customer surveys (i.e., *NPS*, *Politeness*, *Knowledge*, *Satisfaction*, and *Helpfulness*), we eliminate two stores for which we had only post-period observations. We also calculated alternative versions of the *NPS* variable in line with Casas-Arce, Lourenço, and Martínez-Jerez (2017) and estimated all statistical models with these alternative measures, obtaining equivalent results.

the store managers have 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 position, considering the number of positions available and employees she/he has to compete with for that position.<sup>31</sup> On average, in the pre-intervention period, stores held about 25.8 days of sales in inventory. About 13% of store managers had received their compensation late at least once in the pre-period.

Table 1, Panel B reports tests of covariate balance between our control stores 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. We find no other significant differences. In all our analyses we include store fixed effects to control for time-invariant store characteristics (which addresses differences in the stores' average pre-period sales).

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

Table 2 reports the Pearson correlation coefficients calculated across all 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 concern the managing director had that sales may be driven by aggressive sales tactics rather than customer service in some cases. As expected, the measures extracted from the customer surveys are internally consistent, with positive correlations between *NPS*, *Helpfulness*, *Politeness*, *Knowledge*, and *Satisfaction*. 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*). Greater opportunities for promotion (*Promotion Opportunities*) are associated with lower sales and lower incidence of *Failed Audits*, suggesting that sales teams may focus on a broader set of behaviors in line with long-run organizational values. Stores with greater opportunities for promotion have more frequent store manager changes and less experienced managers. Our overall assessment of the

---

<sup>31</sup> This is just a high-level interpretation considering the fact that our measure captures multiple types of promotions for which different sets of store employees could be eligible. For greater detail see footnote 26.



correlations among our predictors is that the risk of collinear relations in the definition of our statistical model is not material.<sup>32</sup>

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

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

Our first set of hypotheses captures our predictions with respect to the effect of the intervention on dimensions of performance connected with the company's four core values. We test these hypotheses using the following difference-in-differences statistical model:

$$Performance_{i,t} = \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Store Manager Change_{i,t} + \beta_4 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon \quad (1)$$

where the dependent variable *Performance* is substituted by each of the dependent variables described at the beginning of section 4.1.1; the indicator variable *Post* equals 1 if week *t* is after the implementation of the 360-degree system, and 0 otherwise; the indicator variable *Treatment* equals 1 if the system is implemented in store *i*, and 0 otherwise;<sup>33</sup> 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;<sup>34</sup> and *Sales Days* is the number of days in the week where the store had at least one sales transaction. To account for unobservable factors associated with each particular store, all our estimations include store fixed effects. Since we observe repeated measures of performance for each store in our sample, standard errors are clustered by store in all our estimations.

#### 4.1.3 Analysis of Organizational Factors Moderating the Effect of the Intervention (H2, H3)

We explore the moderating effect of organizational factors that are likely to influence the response to the introduction of a values-based 360-degree assessment system. In particular, *Hypothesis 2* predicts that the availability of greater opportunities for promotion will accentuate the *evaluative* mechanism of a 360-degree assessment system, thus leading to greater performance improvements in stores with greater

---

<sup>32</sup> We corroborate our conclusion by running several tests of collinearity in the estimation of the coefficients in our model. For example, VIFs are all below 2.

<sup>33</sup> The main effect of *Treatment* does not appear in Equation (1) since it is absorbed by the store fixed effects.

<sup>34</sup> This variable is set to zero for all stores in all pre-intervention period weeks.

promotion prospects. To the extent store members saw the 360-degree system as an evaluative mechanism, the system could give rise to stronger implicit incentives in locations with greater opportunities for promotion.

In testing *Hypothesis 3*, we examine whether the effectiveness of the 360-degree system depended 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 expected the store managers' experience working with top management to be a driver of the system's success.

To examine whether these organizational factors moderated the effect that the 360-degree system had on the outcomes analyzed, we estimated the following model:

$$\begin{aligned}
 Performance_{i,t} = & \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Post_t * Moderator_i \\
 & + \beta_4 Post_t * Treatment_i * Moderator_i + \beta_5 Store Manager Change_{i,t} \\
 & + \beta_6 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon
 \end{aligned} \tag{2}$$

When performing our multivariate tests of *H2*, we substituted *Moderator* with *Promotion Opportunities*, while in our tests of *H3*, *Moderator* was substituted by *Experienced Manager*. Because our moderator variables are defined as time-invariant, store-level characteristics and all our estimations include store fixed effects, the interaction between *Treatment* and *Moderator* is subsumed in both cases.

## 4.2 Results of Statistical Tests

### 4.2.1 Multivariate Tests of the Average Effects of the Intervention (*H1*)

We perform separate OLS estimations of Equation (1) for each of our eight dependent variables of interest. Since Equation (1) is specified using a difference-in-differences approach, our primary coefficient of interest is  $\beta_2$ , which is associated with the interaction term (*Post\*Treatment*). Contrary to expectations, 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). Therefore, our analyses provide no empirical support for *H1*.<sup>35</sup>

---

<sup>35</sup> We note that we find a negative coefficient on *Post* for most of the dependent variables collected through the customer surveys (*Knowledge, Satisfaction, Politeness, and Helpfulness*). We believe this is likely explained by a

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

Descriptive evidence gleaned from interviews indicates that the staff members viewed both the session launching the 360-degree system and the completion of the related surveys positively. They expected the system to yield several beneficial effects including enhancing transparency, motivating employees, 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 their store manager's performance and some respondents mentioned that the purpose of the system was to motivate the store manager and/or help him to grow/improve. Despite the positive perceptions of the system, however, our interviews revealed some potential issues related to the understanding and recall of the core values.<sup>36</sup> Insufficient reinforcement of the core values could explain, in part, the lack of average results for the intervention.

The overall positive reception of the intervention, however, further highlighted the importance of understanding the impact that organizational factors may have on the effectiveness of the introduction of a values-based 360-degree assessment system.

#### *4.2.2 Multivariate Tests of the Moderating Effects of Promotion Opportunities Accentuating the Evaluative Effects of the 360-degree System (H2)*

Table 4 reports our OLS estimation results of Equation (2) where *Moderator* is substituted by *Promotion Opportunities*. The coefficient associated with the moderation effect ( $\beta_4$ ) is positive and

---

change in the company personnel conducting the customer surveys early in the post-period rather than an actual decline in these behaviors in the post-period on average (for instance, it is possible that the new staff provided slightly different answers to clarification questions). Since this change should have affected all stores equally there is no reason to expect that it would affect the estimates for our coefficients of interest to this study (e.g., *Post\*Treatment*).

<sup>36</sup> To contain the risk of contamination between our treated and control stores, participants to the intervention launch event were not allowed to keep a copy of the core values. Additionally, as mentioned in footnote 2x, management did not make any explicit reference to the core values or the 360-degree surveys during the weekly meetings with the stores, as those meetings involved participants from both the treatment and control groups of stores. Consequently, very few of the fourteen interviewees could recall the core values. Those that said they could recall them provided only partial descriptions of the values: an interviewee recalled one core value correctly, and not the others; another interviewee recalled one core value correctly and recited a second core value that was a blend of two actual ones; and a third interviewee listed four core values, none of which were correct but were related to the core value of working hard. Of the two store managers interviewed, one recalled the core values partially, while the other did not recall them but said that he knew what they meant once the interviewer showed them to him.

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 51.8% in sales, 21 percentage points in the *NPS* score, and 5.5 points in the percentage of invoices including bundles, for treated stores relative to control stores, following the introduction of the 360-degree system.<sup>37</sup> In other words, the 360-degree system was associated with more favorable changes in sales, net promoter score, and cross-selling, the greater the stores' opportunities for promotion.<sup>38</sup>

These results are consistent with insights from our interviews, where some of the store team members expressed excitement about how the 360-degree system could help their career aspirations within RETAILER. In one of our conversations with the managing director, 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 to a cashier or store manager position.

Taken together, our analyses suggest that the values-based 360-degree system was most effective when the *evaluative mechanism* of the system was more pronounced. Specifically, *Promotion Opportunities* moderated the effect of introducing the values-based 360-degree system on performance dimensions related to two of the four core values the company was introducing (related to working hard and giving more value to customers), providing partial support for hypothesis 2. 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 values where the opportunities for promotion were higher.

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

---

<sup>37</sup> 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*) = 3.477. 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:  $e^{(0.120*3.477)} - 1 = 0.5178$

<sup>38</sup> We subject all of our statistical analyses to a battery of robustness tests, which we summarize in Appendix 5.

#### 4.2.3 Multivariate Tests of the Moderating Effects of the Store Manager's Experience Accentuating the Developmental Effects of the 360-degree System (H3)

Results of an OLS estimation of Equation (2) where *Moderator* is substituted by *Experienced Manager* are reported in Table 5. Consistent with H3, we find significant moderating effects, whereby the incremental effect of having an experienced manager corresponds to an increase in *Satisfaction* of 0.83 points, in the *Politeness* rating of 0.78 points, and in the *Helpfulness* rating of 1.14 points for treated stores relative to control stores, following the introduction of the 360-degree system.<sup>39</sup> This suggests that, having a store manager with at least six months of experience led to significant marginal intervention effects on measures of performance that were not previously explicitly incentivized but that captured two of the core values promoted by the system related to giving more value to customers and being caring and respectful to others. The fact that the store manager's experience had a stronger impact moderating the 360-degree effects on long-term performance dimensions rather than short-term dimensions already included in explicit incentives makes sense, since the former are the areas where employees were more likely to need to develop new skills.

These results are also consistent with a conversation with the managing director, who noted that, following the intervention, several employees approached her to ask how they could create long-term relationships with their customers (a skill that they realized they needed in order to implement the core value "*We give more value*"). In her words: "They are asking, 'How do we live [the "*We give more value*" core value] 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?'" In this context, our quantitative and qualitative results suggest that the 360-degree system could be more effective where store managers had the knowledge and experience necessary to help their teams embrace the company's values.

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

---

<sup>39</sup> Recall that *Satisfaction*, *Knowledge*, *Politeness* and *Helping* are assessed on a scale from 1 to 5.

## **5. SUPPLEMENTARY ANALYSES**

In this section, we discuss supplementary analyses that we ran to explore two additional conditions, uncovered during our follow-up interviews with store team members and with the managing director, that may have moderated the effectiveness of the values-based 360-degree system and perhaps contributed to the non-significant average effects of the system. Specifically, we explore whether the effects of the 360-degree system were contingent on a store's availability of inventory or delays in salary payments.

While the manager director was generally viewed positively and considered to be supportive, our interviews revealed that several staff members had complaints regarding management exhibiting behaviors that were not in line with the core values highlighted by the 360-degree survey. 29% complained that they did not have enough inventory on hand to satisfy the requests of customers, which appeared to be inconsistent with the principle of giving more value. 29% complained about salaries not being paid on time, which clashed with the core value of being honest and trustworthy. Such issues could potentially lead to demotivation amongst store teams and could, therefore, not only mute the effect of the intervention, but even elicit negative reactions to management's request to further improve performance. We conjecture that store managers feeling that the organization did not hold up their end of the deal might have perceived management's request to take on additional responsibilities related to the RETAILER's core values to be unreasonable or unfair. We, therefore, explore the influence that these factors might have had on the effectiveness of the intervention.

### **5.1 Availability of Inventory**

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 faced a constant trade-off between minimizing the stock on hand and risking that the stores may not always have what the customer wanted or enough items to sell bundled solutions.

We explore whether the availability of inventory moderates the performance effects of the 360-degree system. Since the company systematically placed less inventory in certain stores, we expected that those

stores would experience greater challenges than others to improve their performance in response to the introduction of the 360-degree system. Thus, we estimate Equation (2) substituting the variable *Moderator* with *Low DSI*. DSI is days sales in inventory and is calculated as the ratio of 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 months included in the pre-period.<sup>40</sup> We then partition the distribution of pre-period average DSI into quintiles, given that we identified a discontinuity in DSI at the lowest quintile. *Low DSI* is equal to 1 if the store was in the lowest quintile of the distribution based on days of sales in inventory (DSI) in the pre-period, and 0 otherwise.

Table 6 reports the estimation results when we measure managerial support based on the *Low DSI* indicator. We find that low availability of inventory has a negative incremental effect on treated stores relative to control stores following the intervention, reflected in a marginal *reduction* of invoices including bundles by about 18.5 percentage points ( $\beta_4=-0.185$ ,  $p<0.01$ ). This result is consistent with a less effective intervention 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.

----- Insert Table 6 here -----

## 5.2 Delays in Salary Payments

Occasionally, management retained cash payments related to store employees' compensation for a few days for different reasons (e.g., as it tried to resolve issues such as stock missing from those particular stores).<sup>41</sup> We use *Late Payment* as an additional variable capturing organizational circumstances in which store teams might perceive an inconsistency between the 360-degree system communicating values they are being asked to uphold and observable behaviors by management. This is 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.,

---

<sup>40</sup> Our decision to use pre-period *Low DSI* is informed by the possibility that changes in sales performance subsequent to the implementation of the 360-degree system might impact the DSI measure even in the absence of changes in inventory management policy.

<sup>41</sup> RETAILER had a policy by which staff members were held accountable for any missing items from the inventory, as well as any cash discrepancy in the store. Payment of wages were delayed while management acted to reconcile any differences and investigated possible causes.

stores within a 1-mile radius) at least one time during the pre-period, and zero otherwise. Table 7 documents 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 introduction of the 360-degree system is reflected in *Sales* declining by 32.2%, *NPS* 52.2 percentage points lower, the *Knowledge* rating 0.57 points lower, and the *Politeness* score 0.37 points lower in the post-period.

----- Insert Table 7 here -----

The results of this study were informative to the managing director, who then took steps to further promote and clarify the company's values and to improve support, including paying salaries on time and increasing support to store managers through coaching and systems.

## **6. CONCLUSIONS**

This study documents the results of a field experiment conducted in collaboration with a mobile phone retailer in India. The study explores the performance effects of a values-based 360-degree performance assessment system, introduced in tandem with a formalized vision and core values in half of the retail chain's stores, as a means to improve alignment in an organization already motivating its employees via incentives tied to financial performance. The system was designed to introduce developmental and evaluative mechanisms that would motivate employees to balance short-term results with long-term organizational objectives.

We examine the effects of the introduction of the system, before any feedback was provided, on outcomes that were immediately actionable and largely controllable by employees in their day-to-day work. The intervention yielded interesting results. While do not find any significant performance effects on average, we find that the effectiveness of the system improved when the evaluative and/or developmental mechanisms of the underlying values-based 360-degree system were more accentuated. Specifically, we find that the system was more effective in stores where teams had greater opportunities for promotion. In these stores, the evaluative role of the 360-degree system was enhanced, as employees had implicit incentives to act in accordance to the values reflected in the system in order to be promoted. Also, we find



that the system was more effective where store managers had greater experience related to acting on the core values. In this case, the developmental role was enhanced as store managers could better emphasize, model, and promote behaviors aligned with values among their store teams. Conversely, we find that demotivating factors, such as not having adequate inventory to deliver on the promise of providing greater value to customers, or not being paid on time, led to a deterioration of performance after the intervention, consistent with employees perceiving an imbalance between management's request for greater effort on core-value-related behaviors and not always acting in line with the same values. These latter demotivating factors may contribute to explaining why the overall average results of the system implementation were insignificant.

Our findings contribute to the literature on 360-degree performance appraisal systems by highlighting organizational factors that influence their effectiveness. We contribute to the study of developmental and evaluative mechanisms used to motivate performance, and highlight demotivating factors that could potentially derail the implementation of a values-based 360-degree performance appraisal. Finally, we contribute to the accounting literature examining the use of multiple measures in performance evaluations.

Our results are generalizable to companies considering a 360-degree system in a setting with preexisting high-powered incentives (though being focused on a single organization, our study is subject to the usual concerns regarding generalizability). We acknowledge that our relatively small sample and potentially imperfect proxies for our core values could have hampered our ability to detect significant results in addition to those we document. However, our proxies allowed us to use more objective measures than the ratings traditionally used to test the effects of performance assessments interventions and to capture both potential causal effects and moderated effects of the values-based 360-degree system.

Future research may address follow-up questions, such as exploring how much communication and reinforcement of core values is necessary for a company's values to resonate with—and be remembered by—employees and what are the best ways to do so. We look forward to future research in this area.

## Appendix 1: Performance Appraisal Instrument

<p><b>Core Value #1:</b>  <b>We Gain Control of Our Own Career By Working Hard Every Day and Reaching Out for Support</b></p>
<p>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.</p> <p>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.</p>
<p><i>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.</i></p>
<p>1.1 Do you explain to all promoters and the cashier their targets and the reward program?</p>
<p>1.2 Do you give all promoters and the cashier their daily target every morning?</p>
<p>1.3 Do you ask all promoters and the cashier every day about their target achievement till date?</p>
<p>1.4 Do you remain positive about targets even if there have been some bad days?</p>
<p>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)?</p>
<p>1.6 Do you work hard to help everyone achieve targets even if popular models are out of stock?</p>
<p>1.7 Do you tell all promoters and the cashier to sell old and stuck models?</p>
<p>1.8 Do you sell to the customer when the brand's promoter is not present?</p>
<p>1.9 Do you make sure any problems in the store (PC, printer, AC, lights, sign board problems, etc.) get fixed?</p>
<p>1.10 Do you stay late if a customer walks into the store at the time of store closing?</p>
<p>1.11 Do you make efforts to get a new promoter if a current promoter resigns or is absent for a long time?</p>
<p>1.12 Do you help get stock from other stores when needed?</p>
<p>1.13 Do you make all the promoters and the cashier believe that they can have a successful career at RETAILER?</p>
<p>1.13b. Do you believe that you can have a successful career at RETAILER?</p>

## Appendix 1: Performance Appraisal Instrument (Continuation)

<b>Core Value #2: We Give More Value</b>
<p>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.</p> <p>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.</p>
<p><i>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</i></p>
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?
<p>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)</p> <p>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</p>
2.11 Do you instruct all promoters and the cashier to make only accurate commitments to the customer?
<p>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)</p> <p>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</p>
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: Performance Appraisal Instrument (Continuation)**

<b>Core Value #3: We are Honest and Ethical</b>
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.
<i>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.</i>
<p>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)</p> <p>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</p>
3.2 Do you tell all promoters and the cashier to be trustworthy to customers?
<p>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)</p> <p>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</p>
<p>3.4 Do you report wrong activity against the company to HO?  Please select "Cannot Answer" if there hasn't been any wrong activity.</p>
<p>3.5 Are you honest at the store? (Being honest means not doing any wrong activity)</p> <p>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</p>
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: Performance Appraisal Instrument (Continuation)

<b>Core Value #4: We are Caring and Respectful</b>	
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.	
<i>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.</i>	
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?

## Appendix 2: Organizational Goals and Values at RETAILER

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

### Appendix 3: Definition of Variables Used in This Study

Variable	Definition
Dependent Variables	
<i>Ln(Sales)</i>	Natural logarithm of weekly net sales for store <i>i</i> in week <i>t</i> .
<i>NPS</i>	Net promoter score for store <i>i</i> in week <i>t</i> . 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).
<i>Politeness</i>	Average score awarded by surveyed customers to store <i>i</i> in week <i>t</i> 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.”
<i>Helpfulness</i>	Average score awarded by surveyed customers to store <i>i</i> in week <i>t</i> 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.”
<i>Knowledge</i>	Average score awarded by surveyed customers to store <i>i</i> in week <i>t</i> with respect to the question “Were the people who served you able to explain the features of the products and services in detail?” 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.”
<i>Satisfaction</i>	Average score awarded by surveyed customers to store <i>i</i> in week <i>t</i> with respect to the question “Were you satisfied with your overall experience at RETAILER?” 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.”
<i>% Invoices with Bundles</i>	Percentage of invoices with bundles of products or services for store <i>i</i> in week <i>t</i> .
<i>Failed Audit</i>	Indicator variable equal to 1 if store <i>i</i> does not pass a random audit examining missing inventory or cash in week <i>t</i> , and zero otherwise.

(Appendix 3 continues on the next page)

Appendix 3 – Continued

Explanatory Variables	
<i>Post</i>	Indicator variable equal to 1 if the week/month is after the introduction of the 360-degree survey instrument, and 0 otherwise.
<i>Treatment</i>	Indicator variable equal to 1 if store <i>i</i> is in the treatment group, and 0 otherwise.
<i>Store Manager Change</i>	Indicator variable equal to 1 if the store manager at store <i>i</i> in week <i>t</i> is different than the store manager that was at the store at the time of the intervention, and 0 otherwise. <sup>42</sup>
<i>Sales Days</i>	Number of days where store <i>i</i> had at least one sales transaction in week <i>t</i> .

Moderator Variables	
<i>Promotion Opportunities</i>	Variable calculated as (# stores within a 1-mile radius) / (# employees competing for a potential promotion within a 1-mile radius). The number of employees competing for a promotion is calculated as the median number of employees working in store <i>i</i> plus the median number of employees working at stores within a mile-radius from store <i>i</i> , calculated over the pre-period, minus one (because an employee does not compete with himself). For the purpose of this definition, employees include the store manager, the cashier, and the promoters.
<i>Experienced Manager</i>	Indicator variable equal to 1 if 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.
<i>Low DSI (Pre-Period)</i>	Indicator variable equal to 1 if the average value of monthly days of sales in inventory in the pre-period for the store was in the bottom quintile, and 0 otherwise.
<i>Late Payment</i>	Indicator variable equal to 1 if the store manager at store <i>i</i> was paid at least one time after at least half of the stores located in a one-mile radius from store <i>i</i> in the pre-period, and 0 otherwise.

<sup>42</sup> 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: Follow-up interview questionnaire

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

## Appendix 5: Summary of Robustness Tests

	<i>Overall Intervention (Table 3)</i>	<i>Moderating Effect of Promotion Opportunities (Table 4)</i>	<i>Moderating Effect of Tacit Knowledge (Table 5)</i>	<i>Moderating Effect of Low Availability of Inventory (Table 6)</i>	<i>Moderating Effect of Late Compensation (Table 7)</i>
<b>Main Results</b>	No Significant Effects	Improved <i>Ln(Sales)</i> , <i>NPS</i> , % <i>Invoices with Bundles</i>	Improved <i>Satisfaction</i> , <i>Politeness</i> , <i>Helpfulness</i>	Declined % <i>Invoices with Bundles</i>	Declined <i>Ln(Sales)</i> , <i>NPS</i> , <i>Knowledge</i> , <i>Politeness</i>
<b>Robustness Tests – Are Main Results Significant and Consistent with Main Tests (Yes/No/No-opp)<sup>43</sup></b>					
Using alternative version of NPS	Yes*	Yes	Yes	Yes	Yes
Excluding completely one control store in which a treated store manager became the store manager starting from July 2015 <sup>44</sup>	Yes*	Yes ( <i>Ln(Sales)</i> , <i>NPS</i> ) No (% <i>Invoices with Bundles</i> )	Yes <sup>45</sup>	Yes	Yes ( <i>NPS</i> , <i>Knowledge</i> , <i>Politeness</i> ) No ( <i>Ln(Sales)</i> )
Excluding completely one control store in which a treated store manager became the store manager starting from June 2015 <sup>46</sup>	Yes*	Yes	Yes	Yes	Yes
Excluding completely a store for which we only had two observations in the pre-period for the customer survey-related variables.	Yes*	Yes	Yes	Yes	Yes

(This table continues on the next page)

<sup>43</sup> 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\*Moderator*, where the moderator 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 the Table 3, which were not significant. Any additional results are reported in footnotes.

<sup>44</sup> In our main tests, we had dropped all weeks after the change in the store manager.

<sup>45</sup> Additionally, we find a negative coefficient for the interaction when the DV is % *Invoices with Bundles*, indicating a deterioration of selling complete solutions.

<sup>46</sup> In our main tests, we had dropped all weeks after the change in the store manager.

(Appendix 5 – cont’d)

	<i>Overall Intervention (Table 3)</i>	<i>Moderating Effect of Promotion Opportunities (Table 4)</i>	<i>Moderating Effect of Tacit Knowledge (Table 5)</i>	<i>Moderating Effect of Low Availability of Inventory (Table 6)</i>	<i>Moderating Effect of Late Compensation (Table 7)</i>
<b>Main Results</b>	No Significant Effects	Improved <i>Ln(Sales)</i> , NPS, % <i>Invoices with Bundles</i>	Improved <i>Satisfaction</i> , <i>Politeness</i> , <i>Helpfulness</i>	Declined % <i>Invoices with Bundles</i>	Declined <i>Ln(Sales)</i> , NPS, <i>Knowledge</i> , <i>Politeness</i>
<b>Robustness Tests – Are Main Results Significant and Consistent with Main Tests (Yes/No/No-opp)</b>					
Restrict sample to stores with no change in the store manager	Yes*	Yes	Yes	Yes	Yes (NPS, <i>Knowledge</i> , <i>Politeness</i> ) No ( <i>Ln(Sales)</i> )
Using <i>Open Days</i> instead of <i>Sales Days</i>	Yes*	Yes	Yes	Yes	Yes
Dropping stores with no close stores <sup>47</sup>	N/A	Yes (NPS) No ( <i>Ln(Sales)</i> ), No-opp (% <i>Invoices with Bundles</i> )	N/A	N/A	Yes (NPS) No ( <i>Ln(Sales)</i> )
Including week fixed effects	Yes*	Yes	Yes	Yes <sup>48</sup>	Yes
Including month fixed effects	Yes*	Yes	Yes	Yes	Yes ( <i>Ln(Sales)</i> , NPS, <i>Knowledge</i> ) No-opp ( <i>Politeness</i> )
Removing <i>Store Manager Change</i> from the predictors	Yes*	Yes ( <i>Ln(Sales)</i> , NPS) No (% <i>Invoices with Bundles</i> ) <sup>49</sup>	Yes	Yes	Yes (NPS, <i>Knowledge</i> , <i>Politeness</i> ) No ( <i>Ln(Sales)</i> )

<sup>47</sup> This robustness test is applicable only to the analyses in which the moderator variable depended on the number of close stores (i.e., *Promotion Opportunities* and *Late Payment*).

<sup>48</sup> In addition, in this case, the coefficient for the interaction estimated when *Failed Audits* is the DV is negative, indicating an improvement in the weekly audits’ performance.

<sup>49</sup> Additionally, we find a negative coefficient when the DV is *Failed Audits*, indicating an improvement in the weekly audits’ performance.

## References:

- 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.
- Atwater, L.E., J. F. Brett, and A. C. Charles. 2007. Multisource Feedback: Lessons Learned and Implications for Practice. *Human Resource Management* 46(2): 285-307.
- Atwater, L.E. and J.F. Brett. 2005. Antecedents and Consequences of Reactions to Developmental 360° Feedback. *Journal of Vocational Behavior* 66(3): 532-548.
- Avery, K. 2000. The Effects of 360 Degree Feedback over Time. *ProQuest Dissertations and Theses*.
- 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., & Boudreau, J. W. 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.
- Campbell, D. 2008. Nonfinancial Performance Measures and Promotion-based Incentives. *Journal of Accounting Research* 46(2): 297-332.
- Casas-Arce, P., S. Lourenco, and F. A. Martinez-Jerez. 2017. The Performance Effect of Feedback Frequency and Detail: Evidence from a Field Experiment in Customer Satisfaction. *Journal of Accounting Research* 55(5): 1051-1088.
- Charness, G., and U. Gneezy. 2009. Incentives to Exercise. *Econometrica* 77(3): 909-931.
- Floyd, & List, J. A. (2016). Using Field Experiments in Accounting and Finance. *Journal of Accounting Research*, 54(2), 437-475.
- Gallani, S. 2022. Conduit Incentives: Eliciting Cooperation from Workers Outside of Managers' Control. Evidence from Hospital Hand Washing. Harvard Business School Working Paper.
- 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.
- Gibbs, M., K. A. Merchant, W. A. Van der Stede, and M. E. Vargus. 2004. Determinants and Effects of Subjectivity in Incentives. *The Accounting Review* 79(2): 409-436.
- 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.
- Hall, M. 2008. The Effect of Comprehensive Performance Measurement Systems on Role Clarity, Psychological Empowerment and Managerial Performance. *Accounting, Organizations and Society* 33(2-3): 141-163.
- 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.
- Ittner, C. D., and D. F. Larcker. 1998. Innovations in Performance Measurement: Trends and Research Implications. *Journal of Management Accounting Research* 10: 205-238.
- Ittner, C. D., D. F. Larcker, and M. W. Meyer. 2003. Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard. *The Accounting Review* 78 (3):725-758.
- Ittner, C. D., D.F. Larcker, and T. Randall. 2003. Performance Implications of Strategic Performance Measurement in Financial Services Firms. *Accounting, Organizations and Society* 28(7-8): 715-741.

- Kaplan, R. S., and D. P. Norton. 1992. The Balanced Scorecard - Measures that Drive Performance. *Harvard Business Review*: 71-79.
- 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>.
- 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.
- Loughry, M. L., and H. L. Tosi. 2008. Performance Implications of Peer Monitoring. *Organization Science* 19(6): 876-890.
- Merchant, K. and T. Sandino. 2009. Four Options for Measuring Value Creation. *Journal of Accountancy*. August Issue: 34-37.
- 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.
- Prendergast, C. 1999. The Provision of Incentives in Firms. *Journal of Economic Literature* 37(1): 7-63.
- 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.
- Simons, R. 2010. *Seven strategy questions : A simple approach for better execution*. Boston, MA: Harvard Business Review Press.
- Smith, A. F. R. and Fortunato, V. J. 2008. Factors Influencing Employee Intentions to Provide Honest Upward Feedback Ratings. *Journal of Business and Psychology* 22(3): 191-207.
- Smither, J. W., M. London, N. Vasilopoulos, R. R. Reilly, R. E. Millsap, and N. 1995. An Examination of the Effects of an Upward Feedback Program Over Time. *Personnel Psychology* 48(1): 1-34.
- Smither, J. W., M. London, and R. R. Reilly. 2005. Does Performance Improve Following Multisource Feedback? A Theoretical Model, Meta-analysis, and Review of Empirical Findings. *Personnel Psychology* 58(1): 33-66.
- 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**

<b>Variable</b>	<b>N</b>	<b>Mean</b>	<b>S.D.</b>	<b>P25</b>	<b>P50</b>	<b>P75</b>	<b>Min</b>	<b>Max</b>
<i>Ln(Sales)</i>	694	12.020	1.426	11.212	12.117	13.014	6.745	14.867
<i>NPS</i>	372	0.340	0.384	0.080	0.333	0.571	-1.000	1.000
<i>Politeness</i>	371	3.822	0.559	3.500	3.778	4.000	1.000	5.000
<i>Helpfulness</i>	371	3.433	0.600	3.000	3.250	3.500	1.000	5.000
<i>Knowledge</i>	371	3.721	0.601	3.417	3.688	4.000	1.000	5.000
<i>Satisfaction</i>	372	3.891	0.555	3.583	3.837	4.142	1.000	5.000
<i>% Invoices with Bundles</i>	694	0.263	0.132	0.165	0.252	0.354	0.000	0.714
<i>Failed Audit</i>	79	0.671	0.473	0.000	1.000	1.000	0.000	1.000
<i>Store Manager Change</i>	694	0.105	0.307	0.000	0.000	0.000	0.000	1.000
<i>Promotion Opportunities</i>	672	0.194	0.120	0.098	0.167	0.235	0.091	0.500
<i>Experienced Manager</i>	636	0.758	0.429	1.000	1.000	1.000	0.000	1.000
<i>Average DSI (Pre-Period)</i>	694	25.812	9.287	17.725	23.631	32.766	14.957	50.062
<i>Late Payment</i>	694	0.127	0.333	0.000	0.000	0.000	0.000	1.000
<i>Sales Days</i>	694	6.914	0.428	7.000	7.000	7.000	1.000	7.000

*Notes:* In order to have a consistent length in the pre- and post- periods we have restricted the sample to include 11 weeks in the pre-period and 11 weeks in the post-period. 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. 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 3.

**Table 1, Panel B: Covariate Balance**

<b>Variable</b>	<b>Control Stores</b>		<b>Treatment Stores</b>		<b>Difference</b>
	<b>N</b>	<b>Mean</b>	<b>N</b>	<b>Mean</b>	
<i>Ln(Sales)</i>	176	11.737	176	12.446	0.709***
<i>NPS</i>	40	0.128	43	0.175	0.047
<i>Politeness</i>	40	4.330	42	4.417	0.087
<i>Helpfulness</i>	40	4.258	42	4.202	-0.056
<i>Knowledge</i>	40	4.227	42	4.189	-0.037
<i>Satisfaction</i>	40	4.355	43	4.410	0.055
<i>% Invoices with Bundles</i>	176	0.299	176	0.280	-0.019
<i>Failed Audit</i>	18	0.944	17	0.882	-0.062
<i>Sales Days</i>	176	6.903	176	6.898	-0.006
<i>Promotion Opportunities</i> <sup>^</sup>	15	0.195	16	0.195	0.000
<i>Experienced Manager</i> <sup>^</sup>	13	0.692	16	0.813	0.120
<i>Average DSI (Pre-Period)</i> <sup>^</sup>	16	28.200	16	23.763	-4.437
<i>Late Payment</i> <sup>^</sup>	16	0.125	16	0.125	0.000

*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 3.

<sup>^</sup> We report only one observation per store since these moderator variables are time-invariant.

**Table 2: Correlation Table**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) <i>Ln(Sales)</i>	1.000						
(2) <i>NPS</i>	-0.050	1.000					
(3) <i>Politeness</i>	-0.042	0.596***	1.000				
(4) <i>Helpfulness</i>	-0.073	0.607***	0.901***	1.000			
(5) <i>Knowledge</i>	-0.102	0.574***	0.921***	0.891***	1.000		
(6) <i>Satisfaction</i>	-0.140	0.667***	0.694***	0.808***	0.637***	1.000	
(7) <i>% Invoices with Bundles</i>	0.362**	0.258	0.204	0.129	0.136	0.092	1.000
(8) <i>Failed Audit</i>	0.322**	-0.041	0.330**	0.308*	0.382**	0.162	-0.017
(9) <i>Store Manager Change</i>	-0.262	-0.106	-0.208	-0.093	-0.133	-0.096	-0.247
(10) <i>Promotion Opportunities</i>	-0.265*	0.042	0.030	0.014	0.064	-0.066	-0.096
(11) <i>Experienced Manager</i>	0.101	-0.061	0.005	0.155	0.079	0.127	0.158
(12) <i>Average DSI (Pre-Period)</i>	-0.471***	-0.050	-0.086	-0.055	-0.071	-0.005	-0.380**
(13) <i>Late Payment</i>	-0.425***	-0.009	-0.086	0.004	-0.052	-0.022	-0.324**
(14) <i>Sales Days</i>	0.144	0.103	-0.255	-0.262	-0.230	-0.176	0.282*

	(8)	(9)	(10)	(11)	(12)	(13)
(8) <i>Failed Audit</i>	1.000					
(9) <i>Store Manager Change</i>	-0.309*	1.000				
(10) <i>Promotion Opportunities</i>	-0.433***	0.414***	1.000			
(11) <i>Experienced Manager</i>	0.171	0.023	-0.334**	1.000		
(12) <i>Average DSI (Pre-Period)</i>	-0.147	0.432***	0.271*	-0.321**	1.000	
(13) <i>Late Payment</i>	0.057	0.053	-0.250	0.059	0.141	1.000
(14) <i>Sales Days</i>	-0.187	0.087	-0.191	0.151	-0.146	0.096

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 3.



**Table 3: Effects of the Introduction of the Values-Based 360° Assessment System on Performance**

This table reports results from running the following regression:

$$Performance_{i,t} = \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Store Manager Change_{i,t} + \beta_4 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon$$

Core Values	<i>We Gain Control of Our Career by Working Hard</i>	<i>We Give More Value</i>				<i>We are Honest and Ethical</i>	<i>We are Caring and Respectful</i>	
Outcomes	<i>Ln(Sales)</i>	<i>NPS</i>	<i>Knowledge</i>	<i>Satisfaction</i>	<i>% Invoices with Bundles</i>	<i>Failed Audit</i>	<i>Politeness</i>	<i>Helpfulness</i>
<i>Post</i>	-0.272** (-2.50)	0.257*** (2.79)	-0.603*** (-4.18)	-0.577*** (-3.51)	-0.061*** (-3.00)	-0.313*** (-2.96)	-0.616*** (-4.51)	-1.034*** (-7.04)
<i>Post*Treatment</i>	0.179 (1.50)	0.012 (0.11)	0.085 (0.47)	-0.015 (-0.08)	0.017 (0.60)	0.006 (0.03)	-0.057 (-0.35)	0.099 (0.56)
<i>Store Manager Change</i>	-0.038 (-0.48)	-0.075 (-1.00)	-0.218** (-2.48)	-0.114 (-1.43)	-0.018 (-0.93)	-0.597*** (-3.67)	-0.121* (-2.02)	-0.132* (-1.91)
<i>Sales Days</i>	0.374** (2.51)	-0.056 (-0.51)	-0.223* (-1.93)	-0.158 (-1.28)	0.005 (0.76)	-0.096 (-1.03)	-0.231** (-2.36)	-0.092 (-1.07)
<i>Intercept</i>	9.524*** (9.33)	0.535 (0.72)	5.729*** (7.46)	5.455*** (6.79)	0.255*** (5.65)	1.577** (2.56)	5.945*** (9.32)	4.852*** (8.68)
<i>Store FE?</i>	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	694	372	371	372	694	73	371	371
<i>adj. R<sup>2</sup></i>	0.206	0.069	0.218	0.235	0.121	0.411	0.307	0.526
<i>adj. R<sup>2</sup> (alt. est.)</i>	0.938	0.097	0.223	0.209	0.654	0.383	0.312	0.514

*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<sup>2</sup> is generally underestimated. In the last row we report the R<sup>2</sup> 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<sup>2</sup>. 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. All variables are defined in Appendix 3.

**Table 4: Moderation Effect of Promotion Opportunities on the Performance Effects of the Introduction of the Values-Based 360° Assessment System**

This table reports results from running the following regression:

$$Performance_{i,t} = \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Post_t * Promotion Opportunities_i + \beta_4 Post_t * Treatment_i * Promotion Opportunities_i + \beta_5 Store Manager Change_{i,t} + \beta_6 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon$$

Core Values	<i>We Gain Control of Our Career by Working Hard</i>	<i>We Give More Value</i>				<i>We are Honest and Ethical</i>	<i>We are Caring and Respectful</i>	
Outcomes	<i>Ln(Sales)</i>	<i>NPS</i>	<i>Knowledge</i>	<i>Satisfaction</i>	<i>% Invoices with Bundles</i>	<i>Failed Audit</i>	<i>Politeness</i>	<i>Helpfulness</i>
<i>Post</i>	-0.317*** (-3.96)	0.211** (2.40)	-0.633*** (-4.25)	-0.650*** (-3.93)	-0.055** (-2.56)	-0.324*** (-3.04)	-0.668*** (-4.85)	-1.075*** (-7.06)
<i>Post*Treatment</i>	0.236*** (2.76)	0.076 (0.71)	0.112 (0.60)	0.060 (0.30)	0.014 (0.49)	0.005 (0.03)	-0.011 (-0.07)	0.135 (0.74)
<i>Post*Promotion Opportunities</i>	-1.790*** (-5.82)	-0.724 (-1.35)	0.177 (0.40)	-0.907 (-1.69)	-0.121 (-1.02)	-0.691* (-1.78)	-0.066 (-0.14)	0.094 (0.16)
<i>Post*Treatment* Promotion Opportunities</i>	3.477*** (7.84)	1.749** (2.39)	0.317 (0.22)	1.801 (1.23)	0.459* (1.71)	0.005 (0.00)	0.298 (0.24)	0.156 (0.12)
<i>Store Manager Change</i>	-0.085 (-1.12)	-0.092 (-1.12)	-0.219** (-2.44)	-0.125 (-1.44)	-0.034 (-1.65)	-0.527** (-2.25)	-0.115* (-1.98)	-0.127* (-1.93)
<i>Sales Days</i>	0.345** (2.53)	-0.085 (-0.75)	-0.189 (-1.66)	-0.123 (-0.98)	0.005 (1.00)	-0.100 (-1.07)	-0.199** (-2.07)	-0.063 (-0.78)
<i>Intercept</i>	9.769*** (10.40)	0.751 (0.98)	5.516*** (7.23)	5.239*** (6.39)	0.252*** (7.74)	1.603** (2.60)	5.747*** (9.17)	4.676*** (8.79)
<i>Store FE?</i>	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	672	360	359	360	672	73	359	359
<i>adj. R<sup>2</sup></i>	0.301	0.070	0.219	0.248	0.143	0.401	0.315	0.529
<i>adj. R<sup>2</sup> (alt. est.)</i>	0.946	0.074	0.223	0.222	0.665	0.362	0.32	0.517

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) / (# employees at the store plus employees at the stores in a one-mile radius), where the number of employees is captured at the median level 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. All variables not defined here, are defined in Appendix 3.

**Table 5: Moderation Effect of Store Manager Experience (Exposure to Values) on the Performance Effects of the Introduction of the Values-Based 360° Assessment System**

This table reports results from running the following regression:

$$Performance_{i,t} = \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Post_t * Experienced Manager_i + \beta_4 Post_t * Treatment_i * Experienced Manager_i + \beta_5 Store Manager Change_{i,t} + \beta_6 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon$$

Core Values	<i>We Gain Control of Our Career by Working Hard</i>	<i>We Give More Value</i>				<i>We are Honest and Ethical</i>	<i>We are Caring and Respectful</i>	
Outcomes	<i>Ln(Sales)</i>	<i>NPS</i>	<i>Knowledge</i>	<i>Satisfaction</i>	<i>% Invoices with Bundles</i>	<i>Failed Audit</i>	<i>Politeness</i>	<i>Helpfulness</i>
<i>Post</i>	0.118 (0.82)	0.348*** (3.28)	-0.317 (-1.07)	-0.106 (-0.31)	-0.081 (-1.66)	-0.667*** (-4.56)	-0.220 (-0.86)	-0.602** (-2.06)
<i>Post*Treatment</i>	-0.165 (-0.73)	-0.019 (-0.14)	-0.248 (-0.80)	-0.658* (-1.92)	0.104 (1.37)	0.228 (1.40)	-0.668** (-2.69)	-0.744** (-2.63)
<i>Post*Experienced Manager</i>	-0.451** (-2.50)	-0.045 (-0.26)	-0.464 (-1.48)	-0.608 (-1.70)	0.048 (0.91)	0.494** (2.65)	-0.503* (-1.80)	-0.630* (-1.99)
<i>Post*Treatment*Experienced Manager</i>	0.390 (1.50)	-0.019 (-0.09)	0.550 (1.49)	0.830** (2.11)	-0.129 (-1.60)	-0.328 (-1.40)	0.780** (2.52)	1.140*** (3.39)
<i>Store Manager Change</i>	-0.016 (-0.20)	-0.089 (-1.04)	-0.189* (-2.01)	-0.111 (-1.45)	-0.022 (-1.23)	-0.630*** (-3.56)	-0.100 (-1.67)	-0.107* (-1.85)
<i>Sales Days</i>	0.369** (2.35)	-0.085 (-0.81)	-0.337*** (-3.32)	-0.202 (-1.66)	0.005 (0.63)	-0.052 (-0.83)	-0.279*** (-2.81)	-0.113 (-1.16)
<i>Intercept</i>	9.744*** (9.08)	0.701 (1.00)	6.514*** (9.77)	5.735*** (7.25)	0.263*** (5.33)	1.290*** (3.22)	6.250*** (9.55)	4.986*** (7.76)
<i>Store FE?</i>	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	636	350	349	350	636	66	349	349
<i>adj. R<sup>2</sup></i>	0.235	0.088	0.243	0.247	0.129	0.452	0.320	0.547
<i>adj. R<sup>2</sup> (alt. est.)</i>	0.935	0.123	0.247	0.225	0.674	0.404	0.327	0.538

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. All variables not defined here, are defined in Appendix 3.

**Table 6: Supplemental Analyses: Moderation Effect of Days Sales in Inventory (DSI) on the Performance Effects of the Introduction of the Values-Based 360° Assessment System**

This table reports results from running the following regression:

$$Performance_{i,t} = \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Post_t * Low DSI_i + \beta_4 Post_t * Treatment_i * Low DSI_i + \beta_5 Store Manager Change_{i,t} + \beta_6 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon$$

Core Values	<i>We Gain Control of Our Career by Working Hard</i>	<i>We Give More Value</i>				<i>We are Honest and Ethical</i>	<i>We are Caring and Respectful</i>	
	<i>Ln(Sales)</i>	<i>NPS</i>	<i>Knowledge</i>	<i>Satisfaction</i>	<i>% Invoices with Bundles</i>	<i>Failed Audit</i>	<i>Politeness</i>	<i>Helpfulness</i>
<i>Post</i>	-0.297** (-2.26)	0.278*** (3.33)	-0.527*** (-2.99)	-0.571*** (-2.80)	-0.075*** (-3.22)	-0.327*** (-2.83)	-0.598*** (-3.50)	-1.041*** (-5.53)
<i>Post*Treatment</i>	0.256* (1.77)	0.056 (0.50)	0.075 (0.34)	0.057 (0.24)	0.059* (1.89)	-0.005 (-0.02)	-0.048 (-0.23)	0.123 (0.55)
<i>Post*Low DSI</i>	0.134 (0.74)	-0.077 (-0.29)	-0.304 (-1.48)	-0.013 (-0.05)	0.076*** (3.06)	0.054 (0.19)	-0.072 (-0.33)	0.031 (0.14)
<i>Post*Treatment*Low DSI</i>	-0.336 (-1.63)	-0.156 (-0.50)	0.078 (0.25)	-0.272 (-0.82)	-0.185*** (-4.73)	0.011 (0.03)	-0.025 (-0.09)	-0.092 (-0.31)
<i>Store Manager Change</i>	-0.043 (-0.54)	-0.084 (-1.10)	-0.234** (-2.59)	-0.122 (-1.44)	-0.020 (-1.02)	-0.580*** (-3.16)	-0.125* (-2.04)	-0.133* (-1.80)
<i>Sales Days</i>	0.376** (2.55)	-0.066 (-0.60)	-0.238* (-1.97)	-0.167 (-1.31)	0.006 (0.92)	-0.085 (-0.73)	-0.236** (-2.31)	-0.093 (-1.06)
<i>Intercept</i>	9.516*** (9.44)	0.601 (0.81)	5.831*** (7.25)	5.523*** (6.62)	0.251*** (6.02)	1.499* (1.93)	5.980*** (8.99)	4.862*** (8.45)
<i>Store FE?</i>	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	694	372	371	372	694	73	371	371
<i>adj. R<sup>2</sup></i>	0.212	0.072	0.221	0.236	0.178	0.394	0.304	0.524
<i>adj. R<sup>2</sup> (alt. est.)</i>	0.938	0.099	0.225	0.209	0.677	0.354	0.309	0.511

*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<sup>2</sup> is generally underestimated. In the last row we report the R<sup>2</sup> 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<sup>2</sup>. 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 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 *Highest DSI* not to be estimated, hence we are not reporting a row for these variables and their interaction. All variables not defined here, are defined in Appendix 3.

**Table 7: Supplemental Analyses: Moderation Effect of Late Payments on the Performance Effects of the Introduction of the Values-Based 360° Assessment System**

This table reports results from running the following regression:

$$Performance_{i,t} = \alpha + \beta_1 Post_t + \beta_2 Post_t * Treatment_i + \beta_3 Post_t * Late Payment_i + \beta_4 Post_t * Treatment_i * Late Payment_i + \beta_5 Store Manager Change_{i,t} + \beta_6 Sales Days_{i,t} + \beta_n (Store Fixed Effects) + \varepsilon$$

Core Values	<i>We Gain Control of Our Career by Working Hard</i>	<i>We Give More Value</i>				<i>We are Honest and Ethical</i>	<i>We are Caring and Respectful</i>	
	<i>Ln(Sales)</i>	<i>NPS</i>	<i>Knowledge</i>	<i>Satisfaction</i>	<i>% Invoices with Bundles</i>	<i>Failed Audit</i>	<i>Politeness</i>	<i>Helpfulness</i>
<i>Post</i>	-0.280** (-2.36)	0.265** (2.54)	-0.561*** (-3.46)	-0.510*** (-2.87)	-0.071*** (-3.31)	-0.304** (-2.47)	-0.596*** (-3.83)	-0.987*** (-5.99)
<i>Post*Treatment</i>	0.228* (1.75)	0.080 (0.67)	0.157 (0.83)	-0.012 (-0.06)	0.028 (0.91)	0.030 (0.15)	-0.009 (-0.05)	0.095 (0.48)
<i>Post*Late Payment</i>	0.080 (0.41)	-0.038 (-0.39)	-0.258 (-1.52)	-0.456** (-2.75)	0.080*** (2.81)	-0.042 (-0.25)	-0.124 (-0.85)	-0.320* (-1.75)
<i>Post*Treatment*Late Payment</i>	-0.390* (-1.72)	-0.522*** (-3.87)	-0.570** (-2.43)	-0.088 (-0.40)	-0.082 (-1.18)	-0.184 (-0.40)	-0.372* (-1.78)	-0.006 (-0.02)
<i>Store Manager Change</i>	-0.052 (-0.66)	-0.083 (-1.09)	-0.218** (-2.39)	-0.098 (-1.14)	-0.022 (-1.26)	-0.615*** (-3.59)	-0.122* (-2.01)	-0.120* (-1.80)
<i>Sales Days</i>	0.376** (2.54)	-0.066 (-0.61)	-0.237** (-2.07)	-0.166 (-1.35)	0.004 (0.66)	-0.112 (-1.11)	-0.240** (-2.47)	-0.097 (-1.12)
<i>Intercept</i>	9.517*** (9.41)	0.593 (0.82)	5.815*** (7.67)	5.503*** (6.92)	0.261*** (6.02)	1.689** (2.51)	5.997*** (9.55)	4.881*** (8.77)
<i>Store FE?</i>	YES	YES	YES	YES	YES	YES	YES	YES
<i>N</i>	694	372	371	372	694	73	371	371
<i>adj. R<sup>2</sup></i>	0.213	0.089	0.238	0.249	0.132	0.398	0.313	0.530
<i>adj. R<sup>2</sup> (alt. est.)</i>	0.939	0.116	0.242	0.223	0.659	0.358	0.317	0.518



*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<sup>2</sup> is generally underestimated. In the last row we report the R<sup>2</sup> 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<sup>2</sup>. 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 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. All variables not defined here, are defined in Appendix 3.