© *Academy of Management Journal*
2021, Vol. 64, No. 1, 293–326.
https://doi.org/10.5465/amj.2018.1408

WHEN DOES PAY FOR PERFORMANCE MOTIVATE EMPLOYEE HELPING BEHAVIOR? THE CONTEXTUAL INFLUENCE OF PERFORMANCE SUBJECTIVITY

WEI HE
Nanjing University

SHAO-LONG LI
Wuhan University

JIE FENG
Rutgers University

GUANGLEI ZHANG
Wuhan University of Technology

MICHAEL C. STURMAN
Rutgers University

An extensive body of literature has demonstrated the incentive effect by which pay for performance (PFP) motivates employees' in-role task performance. Nonetheless, scholars have also posited that PFP is likely to demotivate employees' extra-role behaviors. Drawing upon expectancy theory (Vroom, 1964) and the heuristic processing literature (Kahneman, 2011), we examine the relationship between PFP and employee helping behavior. We perform this examination not only by considering the "pay" component (e.g., PFP intensity) but also by scrutinizing the "performance" component; namely, performance subjectivity, which refers to the extent to which the criteria or indicators used to measure employee performance in the performance appraisal system are subjective. Specifically, we propose that PFP has a conditional positive effect (i.e., in the context of high performance subjectivity) on employee helping behavior, and further theorize and test the underlying psychological mechanism by which individual perceived helping–performance expectancy accounts for the interactive effects between PFP and performance subjectivity on employee helping behavior. The empirical results of three studies employing distinctive methodologies provide general support for our hypotheses. Taken together, our research challenges the conventional wisdom that PFP undermines employees' extra-role behaviors by providing new insight into understanding when and why PFP motivates employee helping behavior.

Pay for performance (PFP), broadly defined as "pay that varies with some measure of individual or organizational performance" (Gerhart & Newman, 2020: 715), is a vital component of an organization's compensation strategy (Gerhart & Rynes, 2003; Gerhart, Rynes, & Fulmer, 2009). Recent estimates have suggested that PFP is appealing for employers, as more than 90% of employers have designed PFP systems that link employees' pay levels with their prior or future job performance (Gerhart & Fang, 2014), seeking to motivate employees' performance improvement (Maltarich, Nyberg, Reilly, Abdulsalam, & Martin,

Correspondence concerning this article should be addressed to Shao-Long Li, Economics and Management School, Wuhan University, China. E-mail: tli@whu.edu.cn.

We gratefully thank Associate Editor Anthony Nyberg and three anonymous reviewers for their insightful comments and guidance. We also thank Peter Bamberger and Elena Belogolovsky for their generous sharing of the experimental materials used in the anagram-solving task. This research was supported by the National Natural Science Foundation of China (grant nos. 71822203, 71772073, 71602147, 71772072).

2017; Nyberg, Pieper, & Trevor, 2016). This incentive effect of PFP can be explained by both economic and psychological theories and principles, such as the incentive intensity principle proposed in agency theory (Jensen & Meckling, 1976) and the instrumentality principle proposed in expectancy theory (Vroom, 1964). Accumulative empirical findings have supported this incentive effect by revealing a positive relationship between PFP and employee in-role task performance (e.g., Han, Bartol, & Kim, 2015; Maltarich et al., 2017; Nyberg et al., 2016; Park & Sturman, 2016). Through both incentive and sorting mechanisms (Cadsby, Fei, & Tapon, 2007), PFP increases individual and collective performance at the team, unit, or firm levels (see Byron & Khazanchi, 2012; Garbers & Konradt, 2014; Jenkins, Mitra, Gupta, & Shaw, 1998 for meta-analytic reviews).

Despite the general evidence of PFP's beneficial effects (Gerhart, 2017; Gerhart & Fang, 2014), researchers have also long raised concerns about the overall effectiveness of PFP (individual performance-based PFP in particular) due to potential unintended consequences of such plans (e.g., Deci, Koestner, & Ryan, 1999; Frey & Osterloh, 2012; Kohn, Stewart, & McAdams, 1993; Lawler, 2000; Pfeffer, 1998). These concerns include the idea that PFP plans may foster coworker envy (Bamberger & Belogolovsky, 2017), encourage unethical behavior (John, Loewenstein, & Rick, 2014; Kouchaki, Smith-Crowe, Brief, & Sousa, 2013), reduce intrinsic motivation and creativity (Eisenberger & Cameron, 1996), and encourage short-term thinking (Deci & Ryan, 1985). In short, although there may be strong evidence that PFP facilitates higher in-role performance and the metrics specifically rewarded by a PFP plan (Gerhart, 2017; Shaw & Gupta, 2015), substantial concern remains that PFP plans may ultimately harm behaviors that fall outside of the PFP plan's domain. One exemplar of these out-of-role behaviors or so-called organizational citizenship behaviors (OCBs) is coworker helping, which is defined as a form of "promotive behavior that emphasizes small acts of consideration" (Van Dyne & LePine, 1998: 109). Helping behaviors are cooperative in nature and include both supporting coworkers and preventing the occurrence of coworkers' work-related problems (Podsakoff, MacKenzie, Paine, & Bachrach, 2000). While PFP specifically encourages particular in-role behaviors, the concern is that PFP may simultaneously cause increased competition, decreased collaboration, and ultimately fewer coworker helping behaviors (Arvey & Murphy, 1998; Deming, 1986; Larkin, Pierce, & Gino, 2012). Given that OCBs in general, and helping behaviors in particular, are considered to play such critical roles in organizational success (see Podsakoff, Whiting, Podsakoff, & Blume, 2009; Podsakoff, Ahearne, & MacKenzie, 1997), a pay system that ultimately inhibits such behaviors could have devastating negative implications on organizational effectiveness.

Upon first examination, the concern that PFP may undermine helping behavior seems to be consistent with the extant theoretical underpinnings of PFP. For example, economic theories assume that individuals are economically rational and tend to maximize their self-interests (Milgrom & Roberts, 1992). As such, the best behavioral strategy for maximizing one's economic returns is to increase efforts in accomplishing in-role work assignments that are formally rewarded in organizations' compensation system; however, such efforts may limit the time, energy, and efforts devoted to extra-role behaviors, such as helping other individuals (Organ, Podsakoff, & MacKenzie, 2006). From a psychological perspective, expectancy theory (Lawler, 1971; Vroom, 1964) similarly posits that individuals are demotivated to engage in behaviors that are not linked with desirable outcomes (i.e., low instrumentality perception), particularly when other behaviors (i.e., in-role behaviors) are explicitly rewarded. This perspective is also consistent with social exchange (Blau, 1964) and psychological contract (Rousseau, 1995) theories, which similarly suggest that organizations employing PFP systems may signal an emphasis on short-term oriented, economic-exchange relationships with their employees. Such signals may make employees feel obligated to complete the prescribed work tasks and responsibilities explicitly required in their employment contracts, which are rewarded by the incentive system, rather than engage in behaviors extending beyond these in-role obligations (e.g., helping other individuals).

However, despite the theoretical and logical arguments suggesting that PFP may inhibit helping behavior, empirical evidence to date has yielded mixed results. Certainly, some studies have shown negative relationships between PFP and extra-role behaviors (Bamberger & Belogolovsky, 2017; Bamberger & Levi, 2009); however, other research has shown nonsignificant relationships (Deckop, Mangel, & Cirka, 1999), while other studies have shown conditionally positive relationships (Deckop, Merriman, & Blau, 2004; Du & Choi, 2010). Furthermore, evidence concerning performance measurement calls into question the theoretical premise underlying the potential negative effect of PFP on helping behavior. Specifically, research has suggested that managers may

evaluate employees' task performance in a heuristic way, causing managers to evaluate employees' complex performance outputs by simply grouping or classifying employees into some meaningful categories (Murphy & DeNisi, 2008; Roberson, Galvin, & Charles, 2007). In turn, employees may engage in opportunistic behaviors to maximize their personal gains (Manthei & Sliwka, 2013) or bolster their leaders' evaluations (Bolino, 1999; Murphy & Balzer, 1986). Indeed, evidence suggests that helping or related affiliative behaviors are commonly considered by evaluators as a part of the employee performance domain, even though such behaviors are distinct from in-role behaviors (Borman & Motowidlo, 1993; MacKenzie, Podsakoff, & Fetter, 1993; Podsakoff et al., 2009; Podsakoff et al., 2000; Rotundo & Sackett, 2002). These varying findings highlight the need to deepen theoretical understanding of the effect of PFP on helping, and specifically underscore the need to consider the relevant context and contingencies when exploring the influences of PFP (Nyberg et al., 2016).

In this research, we challenge the conventional wisdom that PFP undermines employee helping behavior by examining the important, yet often-overlooked, influence of performance measurement, which ultimately influences the nature, type, and meaning of the job performance being rewarded by a PFP plan (Bommer, Johnson, Rich, Podsakoff, & MacKenzie, 1995; Muckler & Seven, 1992; Sturman, 2003; Sturman, Cheramie, & Cashen, 2005). We specifically consider the role of performance subjectivity, which we define as an attribute of the performance appraisal method reflecting the extent to which the criteria and indicators used in the evaluation of employee performance are based on (subjective) human judgments (see Gerhart & Rynes, 2003). Using subjective measures in performance appraisal systems is highly prevalent across occupations and organizations (Levin, 2003; Rynes, Gerhart, & Parks, 2005). These subjective measures are usually behavior-based and process-oriented, such as the quality of the work or service, work proactivity and innovativeness, and coworker satisfaction; thus, "they can be used for any type of job" (Gerhart et al., 2009: 256). Drawing upon expectancy theory (Vroom, 1964) and the heuristic processing literature (Kahneman, 2011), we ultimately propose that PFP has a conditional, positive effect on employee helping behavior in the context of performance subjectivity.

Overall, our research makes several important contributions. First, we challenge the view that PFP has a simple direct (and generally presumably a negative) effect on employees' extra-role behaviors by expanding the theoretical horizon to incorporate the contextual contingency effects of performance measurement. We thus contribute to the burgeoning body of literature that has explored PFP's effect on individual extra-role behaviors (e.g., Bamberger & Belogolovsky, 2017; Deckop et al., 2004; Du & Choi, 2010). Second, we advance the PFP literature by emphasizing the importance of the broader context (Nyberg et al., 2016), particularly the performance appraisal context (e.g., how performance is measured and evaluated), in shaping PFP's influences on employees' behavioral outcomes (Gerhart et al., 2009). More importantly, our research provides more nuanced insights into the psychological mechanism (i.e., helping–performance expectancy) that explains how the context of performance measurement moderates PFP's influences on employees' helping behaviors. Finally, by incorporating the heuristic processing literature and demonstrating the helping–performance expectancy mechanism that explains *when* and *why* PFP motivates employees' helping behaviors, our research provides new insights by applying expectancy theory to predict not only the amount but also the direction of individual effort.

THEORETICAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

The Incentive Effect of PFP and Employee Extra-Role Behavior

According to expectancy theory (Vroom, 1964), individual work performance is positively influenced by one's perception of the strength of linkage between work effort and work performance (i.e., expectancy) and the perception of the extent to which work performance is linked with other desirable outcomes (i.e., instrumentality), such as through the PFP system. Despite the extensive empirical evidence regarding PFP's incentive effect on in-role task effort and performance (Gerhart, 2017), researchers have cast doubt on PFP's influence on employee extra-role behaviors.

The argument that PFP will negatively affect extra-role behaviors is largely based on two perspectives. One key perspective is that engaging in extra-role behaviors may thwart the completion of in-role tasks and work assignments (Bergeron, 2007), undermining performance evaluation results and consequently lowering economic returns under a PFP system. This perspective is consistent with original

conceptualizations of OCBs, defined as "individual behavior[s] that [are] discretionary, not directly or explicitly recognized by the formal reward system, and that in the aggregate [promote] the effective functioning of the organization" (Organ, 1988: 4). Expectancy theory would clearly predict that a PFP plan would not motivate OCBs in unrewarded individuals.

The second key perspective is based on cognitive evaluation theory (Deci, 1975), which suggests that extrinsic rewards (e.g., PFP) have a "crowding-out" effect on intrinsic motivation because external controls undermine one's need for autonomy or self-determination (Ryan & Deci, 2000). Based on this crowding-out effect, some researchers have argued that PFP can effectively motivate employee performance only in a context in which work is routine, repetitive, and tedious but not in a context in which work is nonrepetitive, creative, and more intrinsically motivated (Amabile, 1988).

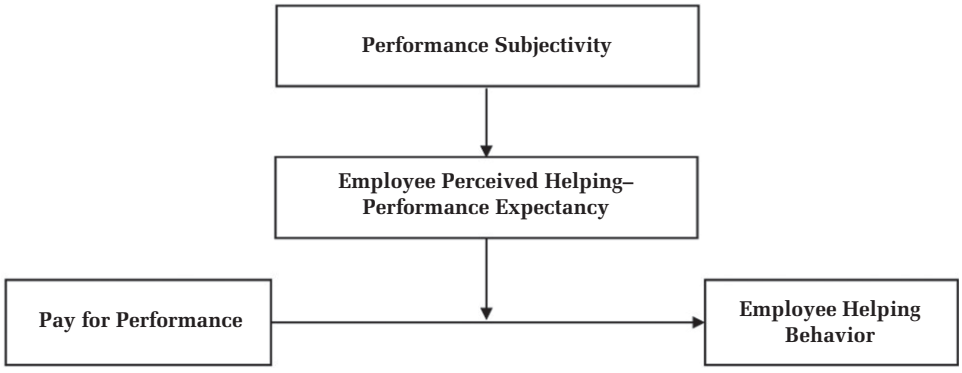
Both perspectives provide a theoretical rationale for why PFP may harm helping behaviors; however, research has suggested that these perspectives are not universally valid. For example, while the crowding-out effect was repeatedly demonstrated in early laboratory studies (Deci et al., 1999), it has received little empirical support in organizational contexts. In fact, extrinsic reward (e.g., PFP) can coexist with or facilitate intrinsic motivation (Gerhart & Fang, 2015). Relatedly, the definition of OCBs has evolved, suggesting that they entail "performance that supports the social and psychological environment in which task performance takes place" (Organ, 1997: 95). The revised definition relaxes the condition that OCBs cannot be rewarded (Podsakoff et al., 2009), calling into question the assumed lack of expectancy and instrumentality associated with pay and OCBs.

These theoretical discrepancies underscore the importance of delving into the specific work context when considering PFP's incentive effect (Nyberg et al., 2016), as the basis for the negative effects of PFP on helping behaviors seems context-sensitive. For example, scholars have posited that a performance-based pay allocation system can motivate employees to help their coworkers in work contexts in which collaboration is emphasized and required (Shaw, 2015; Shaw, Gupta, & Delery, 2002). Indeed, empirical findings have indicated that differentiating team members' pay levels based on their individual performance increases team performance when teams are highly interdependent and, thus, require frequent collaboration (Trevor, Reilly, & Gerhart, 2012). In research focusing on the context of emerging markets as characterized by high collectivism and a socialism-based group orientation, Du and Choi (2010) identified a conditionally positive relationship between PFP and individual helping behavior when the organization had high levels of justice climate or performance appraisal satisfaction as perceived by employees.

Moreover, the helping–performance conflict is more probable when the rewarded performance measure is not associated with (or is even negatively related to) OCBs. This conflict should be most prevalent when performance is results-based and the criteria used to evaluate the performance measurement are highly objective (Rynes et al., 2005). In contrast, in a context in which the performance measurement is behavior-based and process-oriented and the criteria used to evaluate performance are more subjective, helping other individuals may not undermine (and in fact may augment) subjective individual performance assessments (Rotundo & Sackett, 2002). Basing pay on subjective performance ratings may mitigate the potentially dysfunctional effects of PFP by reducing employees' perceptions of risk (Baker, Gibbons, & Murphy, 1994) and lowering "the possibility that employees will focus only on explicitly measured tasks or results at the expense of broader pro-social behaviors, organizational citizenship behaviors, or contextual performance" (Gerhart et al., 2009: 256).

To summarize, the extant empirical findings have suggested that the nature of performance matters in understanding PFP's incentive effect (Gerhart & Fang, 2015). In addition to differentiating employees' performance among in-role routine tasks, in-role creative tasks, and extra-role behaviors, it is important to delve into the performance appraisal context to examine how performance is evaluated and rewarded (Gerhart et al., 2009; Giarratana, Mariani, & Weller, 2018; Nyberg et al., 2016). To gain more nuanced insights regarding the direction and underlying mechanism of the relationship between PFP and employee helping behavior in the context of high or low performance subjectivity, we develop a motivational model (see Figure 1) by drawing on expectancy theory (Vroom, 1964) and focusing on the expectancy component of motivation. Specifically, we argue that the use of a system that offers pay for subjectively evaluated performance enhances the employee's perceived helping–performance expectancy, which we define as the perceived linkage between engaging in helping behaviors and high performance evaluation (see Vroom,

FIGURE 1
The Overall Conceptual Model



1964); this linkage, in turn, explains the interaction between PFP and performance subjectivity on employee helping behavior.

Revisiting the "Performance" Component of PFP

Determining how to define and measure performance is a central issue in conceptualizing PFP. Two dimensions of the "performance" component of PFP should be considered when examining the incentive effect of PFP: (a) the source of performance—that is, whether pay is linked with individual contributions or is collective (e.g., group, unit, and organization) contributions; and (b) the measurability of performance—that is, whether performance measures are objective or subjective (Gerhart et al., 2009).

PFP can be differentiated between individual PFP (e.g., merit pay, sales commissions, bonuses, and cash incentives for individualized contributions [Gerhart & Newman, 2020; Shaw et al., 2002]) and collective PFP (e.g., profit sharing, gainsharing, and stock-related PFP plans [Gerhart & Newman, 2020; Nyberg, Maltarich, Abdulsalam, Essman, & Cragun, 2018]) based on the source or basis of performance. Individual PFP is advantageous for motivating work effort and individual performance, whereas collective PFP helps to maintain relational harmony and collaboration (Gerhart & Rynes, 2003) and provoke team members' prosocial motivation and collaboration (Super, Li, Ishqaidef, & Guthrie, 2016). Bamberger and Levi (2009) demonstrated a positive effect of team performance-based (rather than individual performance-based) pay allocation on team members' helping behaviors. In the present research, we focus on individual PFP (hereafter, PFP) and reexamine its effect on individual helping behavior.

Performance measurability is another vital attribute of a performance appraisal system affecting PFP's incentive effect (Nyberg et al., 2016), and scholars have long been aware of the pros and cons associated with both objective and subjective performance measures (Bommer et al., 1995; Murphy & Cleveland, 1995; Sturman et al., 2005). The traditional economic view considers high performance measurability, such as using objective, results-based measures that are in alignment with organizational goals (Milgrom & Roberts, 1992). It is generally advised for a PFP plan using objective performance measures to provide specific instructions regarding the behaviors and outcomes that lead to rewards, leading to stronger instrumentality as perceived by employees (Gerhart et al., 2009; Murphy & Cleveland, 1995). Fairness or legitimacy problems appear to be another major concern associated with using subjective measures in a PFP system (Rynes et al., 2005) because, in organizational contexts, supervisors tend to have biases (e.g., favorability, severity, leniency, or centrality) in subjectively evaluating employee performance (Marchegiani, Reggiani, & Rizzolli, 2016; Murphy & Cleveland, 1995). Rewarding individual performance without a legitimate basis for what performance should entail may have no incentive effect on individuals, or even detrimental effects on collective performance (Shaw, 2015; Shaw et al., 2002; Trevor et al., 2012).

At the same time, evaluating individual performance based purely on a few objective measures may also generate unintended and dysfunctional consequences (Gerhart et al., 2009) due to the narrow focus on quantitative results and outputs. For example, in a research and development (R&D) context, Giarratana

and colleagues (2018) recently found that rewarding patent inventors' low-quality inventions (inaccurate objective measures) negatively predicted inventors' effort in R&D activities and social interactions with peers. Employees may maximize personal utility by "gaming" the objective performance measures (Baker, 1992), leading to the problem of multitasking tradeoffs (Manthei & Sliwka, 2013) and the misalignment between personal gains and collective interests (i.e., agency problem [Jensen & Meckling, 1976]). Subjective performance measures allow evaluators to consider factors outside the employee's control (Sturman et al., 2005). Thus, subjective measures may facilitate the effectiveness of a PFP system by creating a means-to-ends control system (Baker et al., 1994; Gibbs, Merchant, Van der Stede, & Vargus, 2004; Levin, 2003). As such, subjective measures that emphasize various factors influencing final performance output may have the potential to encourage individuals' extra-role behaviors (Gerhart & Rynes, 2003; Gerhart et al., 2009; Rotundo & Sackett, 2002; Rynes et al., 2005), such as helping coworkers with work-related problems. Next, we elaborate on the crucial role performance subjectivity can play in affecting the relationship between PFP and employee helping behavior.

Performance Subjectivity, PFP, and Employee Helping Behavior

As one exemplar of employees' extra-role behaviors, coworker helping can be both prosocially and extrinsically motivated (Flynn, 2006; Mossholder, Richardson, & Settoon, 2011; Rioux & Penner, 2001). For example, employees help other colleagues as a means to enhance their personal reputation (Bolino, 1999) and accumulate social capital (Mossholder et al., 2011), or because of the collaborative climate or norms in their organizations (Bolino, Turnley, Gilstrap, & Suazo, 2010). Monetary rewards or PFP in particular can also influence employees' motivation toward helping. Specifically, due to the enhanced linkage between employees' actual pay level and individual performance, employees tend to regard any resources within an organization, including good relationships with colleagues, as a means to improve their individual performance and economic returns (Mossholder et al., 2011). Thus, employees may be motivated to engage in helping exchanges with other individuals in order to benefit their personal interests, such as improving their performance evaluations and seeking career advancement (Deckop, Cirka, & Andersson, 2003; De Dreu, 2006; Perlow & Weeks, 2002; Rotundo & Sackett, 2002).

Employees' self-benefiting-motivated helping behaviors in a PFP system can be more prevalent in the work context of high performance subjectivity. When employees' individual performance is less objectively calculated and more subjectively evaluated by their leaders, the line between in-role tasks and extra-role behaviors becomes less obvious (Kamdar, McAllister, & Turban, 2006; Van Dyne, Kamdar, & Joireman, 2008). In other words, there are fewer conflicts or tradeoffs between engaging in in-role tasks and engaging in helping behaviors (Bergeron, 2007). In this context, employees do not decrease their effort in helping others because the instrumental cost is minimal (Eisenhardt, 1989; Jensen & Meckling, 1976). Instead, despite the absence of a key performance indicator (KPI) related to helping behavior on the list, when using subjective measures in a PFP system individuals may internalize coworker helping as part of the individual input valued by the organization and a means to maximize economic returns (Baker et al., 1994; Gibbs et al., 2004; Levin, 2003). Thus, the blurred differentiation between in-role and extra-role behaviors in the context of PFP with subjective measures strengthens employees' instrumental (extrinsic) motivations to help. We therefore propose the following:

Hypothesis 1. Performance subjectivity interacts with PFP to predict employee helping behavior, such that the effect of PFP on employee helping will be more positive when performance subjectivity is high than when it is low.

Psychological Mechanism Toward Helping: Helping–Performance Expectancy

Conventional wisdom based on expectancy theory (Vroom, 1964) suggested that PFP has a negative effect on employee helping behavior as a result of the negative performance expectancy perception that engagement in coworker helping undermines in-role task performance. This perspective, however, is based on the assumption that helping behaviors do not facilitate the acquisition of rewards provided by the system. Given the aforementioned direct and indirect benefits of helping behaviors on the performance outcomes rewarded by a PFP plan, we hypothesized that PFP has a positive effect on helping behavior in the context of high performance subjectivity. The key psychological mechanism underlying

this effect should be exhibited by an explicit belief that there is a link between helping behavior and individual performance appraisal outcome, which we refer to as helping–performance expectancy (see Vroom, 1964).

We argue that a context of high performance subjectivity is associated with employees' enhanced perceptions of helping–performance expectancy. Dual-processing theory (Kahneman, 2011) posits that individuals are likely to engage in a rationality-based, analytic decision-making process when the situation is certain. In contrast, individuals are likely to engage in a heuristic decision-making process when the situation is uncertain. In performance appraisal systems, the use of objective measures typically includes quantitative financial outcomes (e.g., revenue and profits) and quantity of work outputs (e.g., products and services), creating a situation in which employees are certain regarding the behaviors that are expected and rewarded. In this situation, according to expectancy theory (Vroom, 1964), employees may focus exclusively on the ends or results and not the process leading to these ends (Eisenhardt, 1989; Lawler, 1971; Rynes et al., 2005), lowering their beliefs in the linkage between coworker helping and the individual performance evaluated.

In contrast, the use of subjective measures in performance appraisal systems typically involves various qualitative indicators, such as the satisfaction of stakeholders, the quality of products and services, process-oriented work attitudes and behaviors, and adherence to organizational rules, norms, and values. When evaluated according to subjective measures, employees may lack straightforward instructions regarding the behaviors or outcomes that are expected and rewarded, and ultimately engage in heuristic processing to interpret the performance appraisal system. As such, individual behaviors are likely affected by two heuristic biases: (a) representativeness, where individual perceived probability of an event or the frequency of cooccurrences is upwardly biased by the representativeness attributes (e.g., favorability), and (b) availability, where that individual perceived probability of an event or the frequency of cooccurrences is upwardly biased by the ease with which the relevant mental operation of recall or association can be achieved (Tversky & Kahneman, 1974). We argue that, due to these two heuristic biases, employees are more likely to expect helping behavior to boost performance evaluation results in the context of high performance subjectivity.

Specifically, coworker helping is a representative and exemplary OCB that has indirect beneficial consequences for an organization and is thus highly desired by leaders (Podsakoff et al., 2009). In the organizational context, individuals who exhibit frequent helping behaviors are typically referred to as "good soldiers" (Bolino, 1999; Bolino et al., 2010; Organ, 1988) or "extra-milers" (Li, Zhao, Walter, Zhang, & Yu, 2015). These representative and desirable titles make individuals easily associate the occurrence of coworker helping with a favorable performance evaluation outcome (Rotundo & Sackett, 2002; Whiting, Podsakoff, & Pierce, 2008). In addition, although helping behavior is not formally rewarded in the organization's compensation system, its occurrence in the workplace is prevalent as a result of employees' various motives (Rioux & Penner, 2001). Employees are likely to regard helping coworkers as a convenient way to make them look good in front of other individuals, including their leaders who are in charge of performance evaluation, especially when "objective criteria for assessing in-role performance are lacking" (Bolino, 1999: 88). This ease of association of helping behavior with desirable personal returns in the workplace also strengthens individual perceived helping–performance expectancy. Taken together, we argue that individuals' helping–performance expectancy is stronger when subjective measures are employed, versus objective measures, in a PFP system.

Hypothesis 2. Performance subjectivity positively predicts employee perceived helping–performance expectancy.

According to expectancy theory (Lawler, 1971; Vroom, 1964), this enhanced perception of helping–performance expectancy interacts with PFP intensity to motivate individual helping behavior. Expectancy theory posits that individuals exert greater effort while engaging in a particular behavior when they believe that increased effort leads to better performance (i.e., expectancy), which, in turn, is associated with more returns (e.g., pay; instrumentality) that they value (i.e., valence). In other words, individual motivation toward a particular behavior is jointly determined by one's perceptions of expectancy, instrumentality, and valence concerning this behavior (Vroom, 1964). Although linking pay to individual performance that is more subjectively and less objectively measured may generate a weaker incentive effect on in-role task performance (Nyberg et al., 2016), a high PFP intensity

accompanied by subjective measures can increase employees' extra-role helping behaviors because of enhanced perceptions of helping–performance expectancy (resulting from the use of subjective measures) and instrumentality (resulting from the high PFP intensity). Taken together, we propose:

Hypothesis 3. Helping–performance expectancy interacts with PFP to predict employee helping behavior, such that the effect of PFP on employee helping will be more positive when the helping–performance expectancy is high than when it is low.

RESEARCH OVERVIEW

We conducted three studies to test the hypotheses derived from our research model. In Study 1, we collected multisource data from different work occupations in nine organizations. The objective of Study 1 was to test the interactive effects of PFP with performance subjectivity on employee helping behavior (Hypothesis 1). In Study 2, we replicated Study 1's findings and further explored the psychological mechanism of helping–performance expectancy for the moderation effect of performance subjectivity (Hypotheses 2 and 3). We also strengthened the robustness of the empirical results by controlling for more individual, task, and organizational pay-related attributes in Study 2. In Study 3, we conducted a laboratory experiment to establish causality for our hypothesized relationships.

STUDY 1

Participants and Procedure

We collected data from 372 full-time employees in nine Chinese firms. Our research team implemented and collected paper-based surveys following three steps in each organization. First, we obtained information on the organizational structure, pay, and performance appraisal systems and the criteria (objective vs. subjective measures) used in evaluating the individual performance of employees working in different departments and work groups from the human resources (HR) managers. Based on this information, we selected jobs in which either objective measures (e.g., quantity of products) or subjective measures (e.g., managers' assessment of work quality) were used in the performance appraisal system as our targeted research sample. Note that none of these measures explicitly included employees' helping behaviors. Second, we invited employees working in these two job categories to participate in our survey on a voluntary basis. The participants from these two job categories were surveyed independently—that is, at different locations or different time points. Finally, with the help of the HR manager in each firm we collected multisource data. Employees were asked to report the PFP intensity and provide demographic information, and their direct leaders were asked to evaluate the employees' helping behaviors. Table 1 presents the major attributes of our sample firms (i.e., industry and PFP practices) and jobs.

Data from 323 matched leader–employee dyads were collected, resulting in a response rate of 86.83%. In this final sample, the average age was 31 years (SD = 7.76), and the average tenure was 4.59 years (SD = 5.17); 38.39% of the final sample was female, and 87.00% had completed a college education.

Measures

All scale items underwent a back-translation process (Brislin, 1986) to ensure the internal validity of our translated scales.

PFP. Due to the economic reforms in China, PFP systems have been prevalently used in Chinese firms since 1990 (Du & Choi, 2010). Chinese firms that implement PFP systems span nearly all work occupations with varying levels of incentive intensity (i.e., the proportion of floating wage in one's total pay package). To regulate employee work behavior and enhance work motivation in a timely manner, a performance evaluation is typically implemented monthly, and the results of the performance evaluation determine the floating wage that an employee can earn every month.

Chinese firms typically design a PFP system following three steps. First, the amount of the overall (monthly) pay package (level) is decided by a systematic, firm-level job evaluation that assesses and compares the value (e.g., scope, significance, and importance of work responsibilities, workloads, and work environment) of every job in the organization. Second, according to the nature of the work, the jobs are categorized into different series (e.g., administration, sales, R&D, and production), and each category is allocated a different target of the total pay package to be composed of the floating wage, thus setting the incentive intensity. For example, administration jobs might have a target of 0.2 (i.e., 20% percent of floating wage in the total pay package), and sales jobs a target of 0.8. Thus, the nominal floating wage of each job is determined

TABLE 1
Sample Description in Study 1

			High Performance Subjectivity Category			Low Performance Subjectivity Category			
Firm	Industry	Sample Size	Sample Size	Major Occupations	PS Score	Sample Size	Major Occupations	PS Score	Typical PFP Practices
1	Financials	11	4	Bank teller, bookkeeper	4.33	7	Retail client service officer	1.00	Monthly floating wage
2	Consumer services	22	7	Accountant, HR officer	6.33	15	Sales representative	1.67	monthly floating wage, commission
3	Consumer goods	26	22	Accountant, administrative assistant	6.33	4	Process worker	1.67	Monthly floating wage, piece-rate pay, bonus
4	Industrials	23	17	Administrative officer, financial staff	5.33	6	Marketing manager, process worker	2.67	Monthly floating wage, merit pay
5	Industrials	15	2	Accountant, clerk	7.00	13	Sales representative, worker	1.00	Monthly bonus, annual bonus
6	Consumer services	82	5	HR officer, accountant	6.33	77	Business specialist	1.33	Monthly floating wage, annual bonus
7	Consumer services	46	3	Administrative assistant	7.00	43	Purchasing specialist, sales representative	1.00	Commission, annual bonus, monthly floating wage
8	Telecommunications	55	43	Administrative officer, HR officer, accountant	5.67	12	Sales representative, assembly operator	2.33	Monthly floating wage, annual bonus
9	Health care	43	11	Accountant, administrative assistant	5.00	32	Sales representative, purchasing representative	1.00	Monthly bonus, annual bonus, commission

Note: We classified the industry type based on the Industry Classification Benchmark, which was launched by Dow Jones and FTSE in 2005; PS score = performance subjectivity score, as rated by each firm's HR manager.

using the total pay level multiplied by the incentive intensity. Finally, every employee's actual floating wage is calculated by multiplying this nominal floating wage with their monthly performance evaluation index resulting from the performance evaluation result. Thus, the actual floating wage could be more or less than the nominal floating wage depending on the individual's performance appraisal results. In addition to this widely used monthly floating wage, bonus, sales commission, and piece-rate pay are common PFP practices in Chinese firms (see Table 1).

We used one item to evaluate the intensity of PFP involved in each employee's pay package (operationalized as the objective proportion of one's performance-based variable pay in one's total pay). This item was "Please indicate the proportion of your performance-based variable pay (e.g., floating wage, incentives, and bonuses) in your total pay using the following eight categories: (1) 0–5%, (2) 5–15%, (3) 15–30%, (4) 30–50%, (5) 50–70%, (6) 70–85%, (7) 85–95%, and (8) 95–100%." This one-item measure of PFP has been validated in prior research (Du & Choi, 2010).

Performance subjectivity. Following Nyberg et al. (2016), we measured performance subjectivity using a dummy variable. We assigned a value of 1 to this variable for employees who worked in jobs that employed subjective measures in their monthly performance evaluations. In contrast, we assigned a value of 0 for employees who worked in jobs that employed objective measures in their monthly performance evaluations. As shown in Table 1, the sampled employees from the performance subjectivity category were mainly from administration jobs, and the sampled employees from the performance objectivity category were mainly from sales and production jobs.

We sought to strengthen the robustness of our empirical findings by using direct measures of performance subjectivity.[1] A three-item scale was developed following three steps. First, according to the definition of performance subjectivity employed in this research (i.e., the extent to which the criteria and indicators used in evaluating employee performance are subjective), the first author created three initial items measuring the nature (process-oriented or results-oriented) of performance, the measurability (subjectivity or objectivity) of the KPIs, and the overall perception of subjectivity toward the performance appraisal system. All initial items were generated by adapting relevant descriptions from extant performance management and PFP literature (Aguinis, 2009; Gerhart & Rynes, 2003; Gerhart et al., 2009; Murphy & Cleveland, 1995; Rynes et al., 2005). Second, two other authors independently evaluated the content validity of these three items and made modifications. Finally, a sample of 163 MBA students from a large Chinese business school was used to test the reliability ($\alpha = .76$; $\omega^2 = .77$) and validity of these three items; the exploratory factor analysis (EFA) results indicated that the three items loaded on one single factor, with factor loadings that ranged from .74 to .88. These three items were "The criteria or standards used in evaluating my work performance are behavioral and process-oriented rather than results-oriented," "There is no objective KPI that is used in evaluating my work performance," and "In general, my work performance is subjectively evaluated."

We contacted the nine sampled firms again and asked their HR managers to complete a follow-up survey. In this additional data collection, the HR managers were first asked to write down the typical job titles that belong to both subjective performance and objective performance categories, as listed in Table 1. Then, the mangers were asked to rate each job category's performance subjectivity using the three-item scale after shifting the referent. One sample item was "In general, work performance of employees from this job category is subjectively evaluated" (1 = strongly disagree to 7 = strongly agree; $\alpha = .96$ for the nine HR managers' responses on a total of 18 groups of occupation; $\omega = .97$).

Helping behavior. Leaders rated the employees' helping behavior using Podsakoff et al.'s (1997) seven-item scale. Sample items included "This employee is willing to help another employee out if that person falls behind in their work" and "This employee is willing to share their expertise with other members of the team" (1 = strongly disagree and 5 = strongly agree; $\alpha = .87$; $\omega = .87$).

Controls. At the individual level, we controlled for the employees' age, gender, tenure with the organization, and education, which have been found to influence individual helping behavior in prior studies (e.g., Van der Vegt, Bunderson, & Oosterhof, 2006; Van der

[1] We thank the review team for this suggestion.

[2] Coefficient ω is a practical alternative measure of internal consistency to α that has a lower risk of the overestimation or underestimation of reliability and, in many circumstances, outperforms α (e.g., the assumption of scale constancy that commingling or invariant response formats would potentially violate when using α does not underpin ω) (McNeish, 2018).

Vegt, Van de Vliert, & Oosterhof, 2003). In addition, because collective PFP may facilitate coworker helping and collaboration (Nyberg et al., 2018), and since we did not ask participants to report the intensity of collective PFP at the individual level, we added a control for collective PFP at the job level in the previously described follow-up data collection. Specifically, the HR managers were asked to evaluate the incentive intensity of collective PFP in each job category (subjective vs. objective) using a similar one-item measure of individual PFP intensity. At the organizational level, we controlled for the temporal basis of PFP (i.e., a dummy variable that was coded as 1 for firms employing only monthly PFP practices; see Table 1). We also controlled for industry differences. Specifically, according to the Industry Classification Benchmark, our sampled firms can be classified into the following six broad industries: consumer goods, consumer services, industrials, financials, health care, and telecommunications. We created five dummy variables to control for these six industry types.

Results

The means, standard deviations, reliabilities, and correlations are presented in Table 2. Given the nested structure of our data (employees nested in separate companies), we used multilevel path-analytical modeling to test our hypotheses. Model estimation was conducted using Mplus 8.4 (Muthén & Muthén, 2019). To facilitate interpretation of the findings, we group-mean centered Level 1 predicting variables and control variables (except dummies) to obtain an unbiased estimate of the individual-level relationship (Hofmann & Gavin, 1998; Hofmann, Griffin, & Gavin, 2000).

We conducted two path-analytical models, as shown in Table 3, to test our interaction prediction (Hypothesis 1). As shown in Model 1 (i.e., the main effects model), PFP (B = .04, SE = .03, n.s.) and performance subjectivity (B = –.07, SE = .07, n.s.) had no main effects on employee helping behavior. The results of Model 2 (i.e., the interaction effect model) showed that the interaction term between PFP and performance subjectivity was positively related to employee helping behavior (B = .09, SE = .04, $p < .05$). To further explore this interaction effect, we followed Preacher, Curran, and Bauer (2006) and plotted the interaction effect in Figure 2. The simple slope analyses indicated that PFP was positively related to helping behavior among employees working in jobs with subjective performance evaluations (performance subjectivity = 1; simple slope = .10, SE = .04, $p < .05$) but not among employees working in jobs with objective performance evaluations (performance subjectivity = 0; simple slope = .01, SE = .02, n.s.). These results suggest that in the context of performance subjectivity, increasing the PFP intensity from 0–5% to 30–50% (e.g., an increase from category 1 to category 4 in our PFP measure) is associated with an approximate 9% (= (3.80 − 3.50) / 3.50) increase in helping behavior. Thus, Hypothesis 1 is supported.

We conducted additional analyses to test the robustness of our proposed interaction prediction (i.e., Hypothesis 1). We used the three-item measure of performance subjectivity to replace the performance subjectivity dummy and reanalyzed the regression. The results in Model 4 showed that the interaction term, which was created by multiplying the products of two manifest variables (i.e., PFP and performance subjectivity) that were first centered (Aiken & West, 1991), was positively related to employee helping behavior (B = .02, SE = .01, $p < .05$). We plotted the interaction effect in Figure 3. The simple slope analyses showed that PFP was positively related to helping in occupations with high performance subjectivity (performance subjectivity = 2.18; simple slope = .10, SE = .03, $p < .01$) but not among employees working in occupations with low performance subjectivity (performance subjectivity = –2.18; simple slope = –.01, SE = .02, n.s.). These results provide additional support for Hypothesis 1.

Discussion

Through a field study with multisource data from 323 full-time employees, Study 1 provides support for our hypothesized effect that PFP increases employees' helping behaviors in a work context of high performance subjectivity. The results show that, without considering context, PFP seems to be unrelated to employee helping behaviors. When PFP and the performance measurement were considered simultaneously (by examining their interaction), however, PFP was found to lead to helping behaviors when performance is subjectively rated. Interestingly, when performance was objectively rated, PFP had no effect on helping behaviors, thus indicating that PFP may not necessarily harm helping behaviors in many organizational contexts.

Clearly, the implications of this study must be tempered by the nature of its limitations. One critical limitation of this study is that we did not collect data on employees' direct perceptions of performance

TABLE 2
Descriptive Statistics and Correlations of Variables in Study 1

Variables	Mean	*SD*	Correlations														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 Helping behavior	3.64	.59	(.87)														
2 Age	30.58	7.76	.05	—													
3 Gender	.38	.49	.10	.03	—												
4 Tenure	4.59	5.17	−.03	.54	.21	—											
5 Education	2.09	.73	.11	.13	−.09	−.09	—										
6 Collective PFP	2.43	1.23	.10	−.03	−.27	−.36	.05	—									
7 Monthly PFP	.18	.39	.01	.09	−.13	.13	−.04	.11	—								
8 Industry: industrials	.12	.32	.21	−.10	.36	.18	−.06	−.43	−.17	—							
9 Industry: consumer goods	.08	.27	−.03	−.01	−.12	.12	−.04	−.03	.63	−.11	—						
10 Industry: health care	.13	.34	.00	.01	.07	.07	−.07	−.22	−.19	−.14	−.12	—					
11 Industry: consumer services	.46	.50	.14	.07	−.25	−.37	.12	.75	−.09	−.34	−.28	−.37	—				
12 Industry: telecommunications	.17	.38	−.35	−.04	.07	.19	−.07	−.45	−.21	−.17	−.13	−.18	−.42	—			
13 PS dummy	.35	.48	−.19	−.09	.00	.21	−.00	−.70	.20	.11	.31	−.08	−.49	.41	—		
14 PS score	2.88	2.18	−.19	−.09	−.01	.24	−.03	−.69	.26	.11	.37	−.15	−.47	.43	.98	(.96)	
15 PFP	3.03	1.55	.22	.08	.09	.04	.00	.19	.00	.16	−.03	−.03	.17	−.31	−.22	−.22	—

Notes: $n = 323$ at the individual level; $n = 9$ at the firm level. Correlations greater than ± .11 are statistically significant at $p < .05$. Coefficient α estimates of reliability are in parentheses on the diagonal. PFP = pay for performance. PS = performance subjectivity.

TABLE 3
Multilevel Path Analysis Results in Study 1

Variables	Helping Behavior			
	Model 1	Model 2	Model 3	Model 4
Intercept	3.76*** (.13)	3.76*** (.13)	3.75*** (.14)	3.75*** (.14)
Age	−.01* (.00)	−.01† (.01)	−.01† (.01)	−.01 (.01)
Gender	.07 (.04)	.07 (.05)	.08† (.04)	.08† (.05)
Tenure	.01* (.00)	.01† (.00)	.01* (.00)	.00 (.00)
Education	.07† (.03)	.07† (.04)	.06† (.03)	.07† (.04)
Collective PFP			.12*** (.03)	.14** (.05)
Monthly PFP	−.05 (.16)	−.06 (.16)	−.07 (.15)	−.07 (.15)
Industry: industrials	.20 (.19)	.19 (.19)	.16 (.17)	.16 (.17)
Industry: consumer goods	−.08** (.03)	−.09*** (.03)	−.11*** (.00)	−.11*** (.00)
Industry: health care	−.13 (.16)	−.14 (.16)	−.15 (.15)	−.15 (.15)
Industry: consumer services	−.00† (.00)	−.00 (.00)	.00 (.00)	−.00 (.00)
Industry: telecommunications	−.55** (.19)	−.57** (.19)	−.61*** (.15)	−.61*** (.15)
PS dummy	−.07 (.07)	−.06 (.06)		
PS score			.03 (.02)	.05† (.02)
PFP	.04 (.03)	.01 (.02)	.04 (.03)	.04* (.02)
PFP × PS dummy		.09* (.04)		
PFP × PS score				.02* (.01)
Residual variance (within-firms)	.281*** (.03)	.277*** (.03)	.279*** (.03)	.274*** (.03)
Residual variance (between-firms)	.008 (.01)	.008 (.01)	.008 (.01)	.008 (.01)

Notes. $n = 323$ at the individual level; $n = 9$ at the firm level. We used the mean imputation method for 10 missing data points in the analyses. To ensure robustness, we reanalyzed our data by removing these 10 cases and found that these results remained unaffected. In Models 1 and 2, we did not control for collective PFP and used job dummy as the proxy of performance subjectivity (named PS dummy in these models). In Models 3 and 4, performance subjectivity was measured by the firm's HR manager in the additional data collection (named PS score in these models). All coefficients displayed are unstandardized. Standard errors are in parentheses.

† $p \leq .10$
* $p \leq .05$
** $p \leq .01$
*** $p \leq .001$

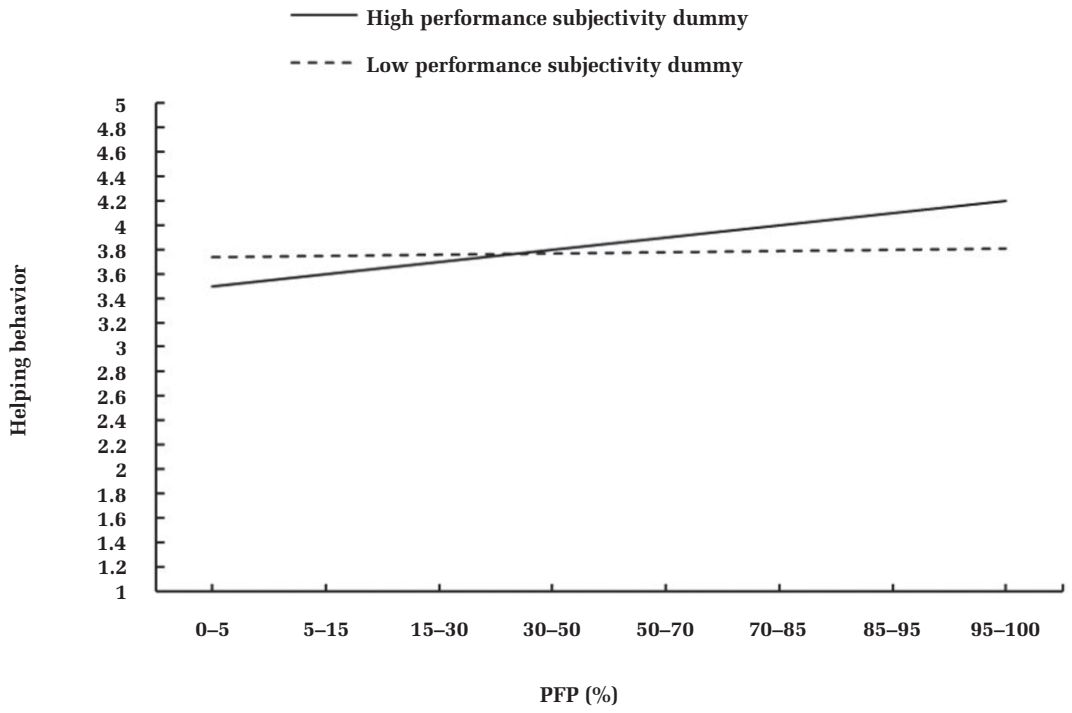
subjectivity. While the use of pay system information (provided by the organization) yields important insights into the effects of the PFP system on employee outcomes (Park & Sturman, 2016), it does not allow us to directly test the mechanisms by which we predict that employee perceptions of pay systems influence their behaviors. Although our findings support our arguments regarding how PFP and the context jointly influence employee helping behaviors, the observed effects of PFP and performance subjectivity may be affected by unmeasured work-related factors, such as task interdependence, indicating the possibility of alternative explanations for our findings. Hence, we conducted Study 2.

STUDY 2

Participants and Procedure

We collected data from six Chinese firms (all different than those used in Study 1). All surveys were implemented via an online survey platform operated by a survey company. With the support of the chief executive officer (CEO) of each firm, we distributed the survey introduction—which explained the purpose, content, and format of the survey—to employees across all jobs. The recruited participants who agreed to enroll in two time-lagged surveys were rewarded 50 Chinese yuan (approximately 7 U.S. dollars). At Time 1, 199 employees from these six firms completed the survey measuring PFP, perceived performance subjectivity, helping–performance expectancy, and the controls. After two weeks, 176 of these 199 employees completed the Time 2 survey measuring individual helping behavior and social desirability measures, resulting in an 88% response rate. There were no significant differences in age, gender, tenure, and education between the Time 1 sample and the Time 2 sample. All data were collected online and anonymously, and the two waves of data were matched by participants' IDs recognized by the online survey platform.

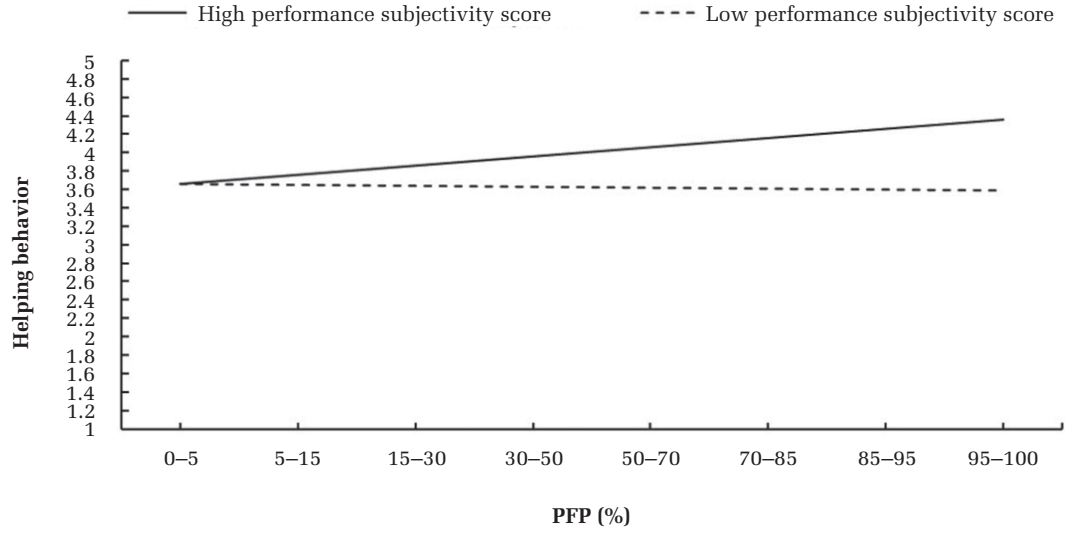
FIGURE 2
The Interactive Effects of PFP with Performance Subjectivity Dummy on Employee Helping Behavior in Study 1



Note: High performance subjectivity dummy = 1; low performance subjectivity dummy = 0.

In this final sample, the average age was 35 years (*SD* = 5.63) and the average tenure was 7.78 years (*SD* = 4.38); 44.9% of the final sample was female, and 97.7% had completed college education. The participants worked in various job categories, including administration (45.5%), R&D (25.6%), sales (17.0%), and production and operation (11.9%).

FIGURE 3
The Interactive Effects of PFP with Performance Subjectivity Score on Employee Helping Behavior in Study 1



Measures

PFP. As a precondition for our data collection, all six sampled firms' PFP systems involved individual monthly performance evaluation and incentives. Moreover, none of these firms' performance evaluation indicators explicitly included employees' helping behaviors. The measure of PFP was improved in several aspects compared with that in Study 1. First, we introduced the definition of individual PFP and provided several examples of individual incentives to participants. Then, we asked employees to recall their monthly pay packages and identified the individual incentives that fit with the definition of individual PFP. Finally, we used one item to evaluate the intensity of individual PFP involved in each employee's monthly pay package. This item was "Please indicate the proportion of your individual performance-based variable pay (e.g., actual floating wage and monthly bonus) in your total monthly pay using the following eight categories: (1) 0–5%, (2) 5–15%, (3) 15–30%, (4) 30–50%, (5) 50–70%, (6) 70–85%, (7) 85–95%, and (8) 95–100%." In general, the improved measure of PFP specified the basis (individual performance), components, and time frame (monthly) of PFP.

Performance subjectivity. We used the same three-item scale of performance subjectivity that was developed and validated in Study 1. The subject rated these three items using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree) with acceptable reliability ($\alpha = .77$; $\omega = .78$).

Helping–performance expectancy. Drawing upon prior research (e.g., Colquitt & Simmering, 1998) measuring expectancy perception (the belief of a linkage between work effort and performance[Vroom, 1964]), we modified the two-item scale of expectancy developed by Lawler and Suttle (1973) to measure helping–performance expectancy. Specifically, we asked employees to indicate how true it was that helping effort led to a higher performance evaluation, using the following two items: "Trying hard to help others increases performance evaluations" and "More helping behavior is associated with better performance evaluations" (1 = strongly disagree to 7 = strongly agree; $\alpha = .83$; $\omega = .83$).

Helping behavior. The participants reported their helping behaviors using the same 7-item scale (Podsakoff et al., 1997) employed in Study 1. These seven items were measured using a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree) and showed good reliability in this study ($\alpha = .77$; $\omega = .77$).

Controls. At the individual level, we controlled for age, gender, tenure with the organization, and education, as indicated in Study 1. In addition, we added individual difference controls of job level (1 = bottom-level employee, 2 = frontline supervisor, 3 = mid-level manager, and 4 = executive manager), job type (three dummies created for four categorical occupations: R&D occupations, sales occupations, production and operation occupations, and administrative occupations), internal pay level (one single item from Blau [1994]: "Compared to relevant employees in my own department or team, may pay level is . . ."; 1 = a lot less and 5 = a lot more), relative performance (one single item: "Compared to relevant employees in my own department or team, my task performance level is . . ."; 1 = much better and 5 = much worse), and risk aversion (four-item scale from Gupta and Govindarajan [1984]; 1 = strongly disagree to 7 = strongly agree; $\alpha = .84$; $\omega = .84$) because prior research has found that these individual differences affect how employees react to PFP (Deckop et al., 2004; Fulmer & Shaw, 2018). We also controlled for individual social desirability using a 10-item scale (1 = strongly disagree to 7 = strongly agree; $\alpha = .82$; $\omega = .83$) developed by Kovacic, Galic, and Jerneic (2014), given that coworker helping was self-reported by the employees.

Because the work context of task interdependence was found to affect employees' reactions to performance-based pay differentials (Shaw, 2015; Shaw et al., 2002; Trevor et al., 2012), we controlled for employees' perceived task interdependence using seven items adapted from Pearce and Gregersen (1991). A sample item was "I work closely with others in doing my work" (1 = strongly disagree to 7 = strongly agree; $\alpha = .80$; $\omega = .81$). For the pay-related information, we controlled for collective PFP (monthly based) using the same one-item measure employed in the robustness analysis of Study 1. Moreover, we controlled for employee perceived pay system communication using the four-item scale developed by Shaw and Gupta (2007). A sample item was "the pay system is clearly communicated to me" (1 = strongly disagree to 7 = strongly agree; $\alpha = .86$; $\omega = .86$). Finally, given that our six sample firms belonged to three different industries (consumer goods, industrials, and technology), we created two industry

dummies to control for industry differences at the firm level, as employed in Study 1.[3]

Results

The means, standard deviations, reliabilities, and correlations of Study 2 are presented in Table 4. We conducted a set of single-level confirmatory factor analyses using Mplus 8.4 (Muthén & Muthén, 2019) to test the discriminant validity among the three focal latent variables in our research model (i.e., performance subjectivity, helping–performance expectancy, and helping behavior). The results showed that our hypothesized three-factor model demonstrated a satisfactory fit to the data ($\chi^2 = 94.67$, $df = 51$, $p < .001$, CFI = .92, TLI = .90, RMSEA = .07, SRMR = .06). In addition, the decreased fit values and χ^2 difference tests suggested that the three-factor model was superior to the other models. For example, (a) in the two-factor model (grouping performance subjectivity and helping–performance expectancy), CFI = .65, TLI = .56, RMSEA = .15, SRMR = .14; $\Delta\chi^2 = 159.96$, $\Delta df = 2$, $p < .001$; and (b) in the two-factor model (grouping helping–performance expectancy and helping behavior), CFI = .79, TLI = .73, RMSEA = .11, SRMR = .08; $\Delta\chi^2 = 80.43$, $\Delta df = 2$, $p < .001$. Overall, the results support the discriminant validity of our focal constructs.

Hypothesis testing. Since the data were nested, we used multilevel path-analytical modeling to test our hypotheses. We first conducted a path-analytical model to show the main effects of PFP and performance subjectivity on individual helping behavior. As shown in Model 1 of Table 5, PFP ($B = .08$, $SE = .05$, n.s.) and performance subjectivity ($B = .00$, $SE = .01$, n.s.) were not significantly related to employee helping behavior. Second, we added the interaction term of PFP with performance subjectivity into the path-analytical model to test Hypothesis 1. As shown in Model 2, the interaction term, which was created by multiplying the products of two manifest variables (i.e., PFP and performance subjectivity) that were first centered (Aiken & West, 1991), was significantly and positively related to helping behavior ($B = .07$, $SE = .02$, $p < .001$). We plotted the interactive effects in Figure 4. Simple slope analyses showed that PFP was positively related to helping behavior among employees working in occupations of high performance subjectivity (performance subjectivity = 1.34; simple slope = .18, $SE = .08$, $p < .05$) but not among employees working in occupations of low performance subjectivity (performance subjectivity = –1.34; simple slope = –.01, $SE = .04$, n.s.). These results suggest that in the context of high performance subjectivity, increasing the PFP intensity from 0–5% to 30–50% (e.g., an increase from category 1 to category 4 in our PFP measure) is associated with an approximate 11% (= (5.60 − 5.06) / 5.06) increase in helping behavior. Thus, Hypothesis 1 is again supported.

Hypothesis 2 predicted a direct effect of performance subjectivity on helping–performance expectancy. As shown in Model 4a of Table 5, performance subjectivity was positively related to employee perceived helping–performance expectancy ($B = .11$, $SE = .04$, $p < .01$). Thus, Hypothesis 2 is supported.

Hypothesis 3 suggested that PFP interacted with helping–performance expectancy to affect employee helping behavior. Consistent with our prediction, the interaction term, which was created by multiplying the products of two manifest variables (i.e., PFP and helping–performance expectancy) that were first centered (Aiken & West, 1991), was significantly and positively related to helping behavior ($B = .07$, $SE = .03$, $p < .05$; Model 4b). The simple slope analyses (see Figure 5) showed that PFP was positively related to helping behavior among employees who perceived higher levels of helping–performance expectancy (helping–performance expectancy = 1.16; simple slope = .17, $SE = .08$, $p < .05$) but not among employees who perceived lower levels of helping–performance expectancy (helping–performance expectancy = –1.16; simple slope = –.00, $SE = .06$, n.s.). Thus, Hypothesis 3 is supported.

Discussion

Through a field study collecting multisource data from 176 full-time employees, Study 2 replicated Study 1's findings. That is, PFP alone was not related to helping behaviors, but when considered in moderation with performance measurement subjectivity, PFP was

[3] For robustness, as well as to rule out the concern of potential *p*-hacking (Sturman, Sturman, & Sturman, accepted), we reanalyzed our data by removing these controls. The results suggested that although the interaction term of PFP with performance subjectivity on helping behavior was marginally significant ($B = .04$, $SE = .03$, $p < .10$), the simple slope of PPF predicting helping in the condition of high performance subjectivity was significantly positive (simple slope = .16, $p < .01$). The interaction term of PFP with helping–performance expectancy on helping behavior was significant ($B = .06$, $SE = .03$, $p < .05$), and the simple slope of PFP predicting helping among employees who perceived a high level of helping–performance expectancy was significantly positive (simple slope = .12, $p < .01$). These results suggest that our hypotheses were largely supported without adding any control variable in this study.

TABLE. 4
Descriptive Statistics and Correlations of Variables in Study 2

Variable	Mean	*SD*	Correlations																	
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1 Helping behavior	5.47	.67	(.77)																		
2 Age	34.19	5.63	.01	—																	
3 Gender	.55	.50	.09	.12	—																
4 Tenure	7.78	4.38	.02	.68	.11	—															
5 Education	2.09	.37	−.03	−.09	.02	−.16	—														
6 Job—R&D	.26	.44	.07	−.05	.01	.00	.08	—													
7 Job—sales	.17	.38	−.02	.20	.14	.19	−.02	−.27	—												
8 Job—production	.12	.33	.01	−.02	.12	−.07	−.13	−.22	−.17	—											
9 Job level	1.84	.85	.19	.21	.30	.11	.14	−.16	.09	.03	—										
10 Pay system communication	5.52	.93	.33	−.02	−.04	−.06	.05	−.00	−.05	−.03	.17	(.86)									
11 Internal pay level	3.61	.63	.18	−.05	.01	−.06	.02	−.03	.06	.00	.28	.39	—								
12 Relative performance	3.91	.56	.23	.11	.04	.05	.07	−.02	.05	.03	.28	.37	.45	—							
13 Risk aversion	4.14	1.25	−.10	.03	.04	−.01	−.08	−.06	−.13	−.06	−.09	.02	−.00	−.11	(.84)						
14 Task interdependence	5.06	.82	.43	.07	.05	.08	.05	.02	.02	−.10	.23	.47	.36	.39	.22	(.80)					
15 Social desirability	3.10	1.91	.19	−.05	.06	−.04	−.04	.08	.00	.03	.04	−.01	−.01	−.09	−.08	−.08	(.82)				
16 Collective PFP	2.48	1.12	.08	.14	.07	.16	.11	.01	.27	−.02	.25	.01	.09	.16	.08	.19	.04	—			
17 Performance subjectivity	3.98	1.34	.01	.08	−.06	.01	.02	−.10	−.20	−.17	.01	.01	.05	−.04	.23	.15	−.09	.01	(.77)		
18 HPE	5.06	1.16	.41	.11	−.00	−.05	.06	−.05	.02	.10	.22	.40	.28	.46	−.01	.47	−.03	.19	.13	(.83)	
19 PFP	2.73	1.09	.16	.05	.05	.06	.00	−.08	.32	.11	.19	.08	.25	.19	−.08	.10	.13	.62	−.00	.19	—

Notes: $n = 176$ at the individual level; $n = 6$ at the firm level. Correlations at or greater than ± .15 are statistically significant at $p < .05$. HPE indicates helping–performance expectancy. Reliability coefficients are displayed within parentheses on the diagonal. Correlations for two industry dummy variables with the other variables are not presented due to space considerations, but are available from the corresponding author.

TABLE 5
Multilevel Path Analysis Results in Study 2

	Testing Hypothesis 1		Testing Hypotheses 2–3		
	Model 1	Model 2	Model 3	Model 4a	Model 4b
Variables	Helping behavior		Helping behavior	DV = HPE	DV = Helping behavior
Intercept	5.36*** (.03)	5.34*** (.02)	5.37*** (.03)	4.83*** (.08)	5.35*** (.03)
Age	.00 (.01)	.00 (.01)	−.01 (.01)	.04* (.02)	−.01 (.01)
Gender	.07 (.05)	.08 (.05)	.09** (.03)	−.11 (.11)	.12*** (.04)
Tenure	−.00 (.02)	−.00 (.02)	.01 (.02)	−.07** (.02)	.01 (.02)
Education	−.14** (.05)	−.13*** (.04)	−.14*** (.04)	−.02 (.07)	−.15*** (.03)
Job: R&D	.05 (.04)	.03 (.04)	.04 (.06)	.05 (.12)	.05 (.06)
Job: sales	−.11 (.10)	−.07 (.10)	−.14 (.09)	.13 (.21)	−.15† (.08)
Job: production	.05 (.11)	.09 (.11)	−.06 (.13)	.55*** (.10)	−.07 (.16)
Job level	.05 (.04)	.05† (.03)	.04 (.03)	.05 (.07)	.04 (.03)
Pay system communication	.10** (.03)	.10*** (.03)	.06* (.03)	.22† (.13)	.05 (.04)
Internal pay level	−.09 (.06)	−.09† (.06)	−.07 (.05)	−.09 (.06)	−.07 (.06)
Relative performance	.02 (.08)	.01 (.08)	−.07 (.08)	.53*** (.13)	−.07 (.09)
Risk aversion	−.09** (.03)	−.09** (.03)	−.08* (.04)	−.06† (.04)	−.08* (.04)
Task interdependence	.34*** (.06)	.36*** (.06)	.27*** (.05)	.44*** (.11)	.27*** (.04)
Social desirability	.06*** (.02)	.06*** (.02)	.06*** (.01)	.01 (.03)	.06*** (.02)
Collective PFP	−.04 (.05)	−.05 (.05)	−.06 (.03)	.09 (.10)	−.06 (.04)
PS	.00 (.01)	.02 (.01)	−.02 (.01)	.11** (.04)	−.02 (.01)
HPE			.17*** (.05)		.19** (.07)
PFP	.08 (.05)	.08 (.05)	.08 (.05)	.02 (.07)	.08 (.06)
PFP × PS		.07*** (.02)			
PFP × HPE					.07* (.03)
Residual variance (within-firm)	.323*** (.04)	.316*** (.04)	.297*** (.03)	.800*** (.05)	.290*** (.02)
Residual variance (between-firm)	.000 (.01)	.000 (.01)	.006 (.03)	.011 (.05)	.005 (.03)

Notes: $n = 176$ at the individual level; $n = 6$ at the firm level. Coefficients for two firm dummy variables included in each calculation are not presented due to space considerations, but are available from the corresponding author. All coefficients displayed are unstandardized. Standard errors are in parentheses. Model 4a and 4b were estimated simultaneously in the same calculation.

† $p \leq .10$
* $p \leq .05$
** $p \leq .01$
*** $p \leq .001$

positively related to employee helping when job performance was perceived as subjectively rated. In addition, Study 2 demonstrated employee perceived helping–performance expectancy as an underlying psychological mechanism accounting for the moderating effect of performance subjectivity. Although we consider the findings of Study 2 more robust than the findings of Study 1 because of our improved measures of the PFP constructs and enhanced control of individual differences, task attributes, and pay-related information, the nature of the data still does not permit the testing of causal conclusions. To address this limitation, we subsequently conducted a laboratory experiment.

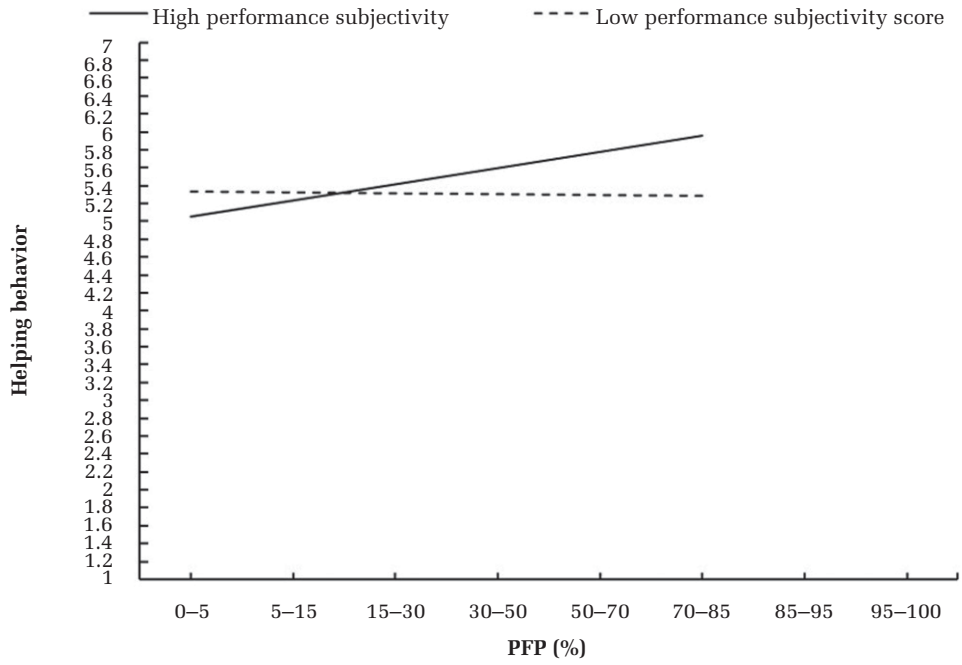
STUDY 3

Participants and Procedure

A total of 168 undergraduate students in a Chinese university were recruited for the experiment via an online advertisement. The participants were informed that they would participate in a simulated business work task through which they could earn monetary rewards. On the experiment day, 160 participants attended, and the participants were randomly grouped into 40 four-person teams to create a social interaction context in which coworker helping was likely to occur. Each team was also randomly assigned a team leader (confederate) to facilitate the process of the experiment and implement performance evaluations and rewards. These 40 confederates, who were senior undergraduates enrolled in the compensation management course taught by the first author, engaged in a pilot experiment to learn the experimental task (as explained below). These confederates were also trained by the first author regarding their responsibilities throughout the experiment.

Before the experiment started, the participants were first gathered in a large classroom to learn the

FIGURE 4
The Interactive Effects of PFP with Performance Subjectivity on Employee Helping Behavior in Study 2

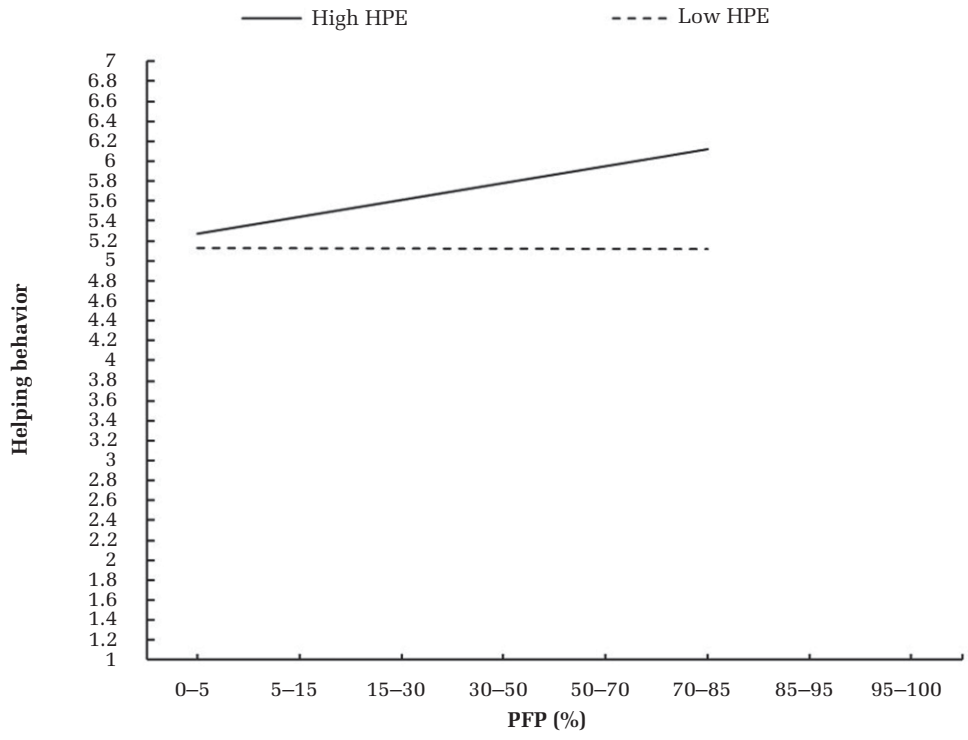


Note: PFP intensity ranged from 1 (0–5%) to 6 (70–85%) in Study 2.

experimental task through an oral presentation provided by the first author, along with a paper-based introduction. The participants were informed that they would play the role of a management consultant for a consulting company and would compete for a consulting project by writing a creative English essay using as many words, resolved from an anagram-solving exercise, as possible. The participants were also informed that the division of teams was meant to facilitate the grading of individual performance in a timely manner, that there was no between-team competition, and that they would receive no monetary reward based on their teams' collective performance. We finally informed the participants of the responsibility of their team leaders, who would provide them with detailed information regarding performance evaluation and monetary rewards and monitor their behaviors throughout the task following the competition rules. After learning about the experimental task and procedure, eight participants, including all four participants from one team, withdrew (e.g., due to personal reasons). Three teams' data were excluded from our analyses because the leaders (confederates) of these teams reported significant violations of the competition rules (e.g., alliance between team members to resist the forced distribution of performance ratings) among the team members. As a result, the remaining 140 participants (M_{age} = 20, 36.4% male) were from 32 four-person teams and four three-person teams. According to our prior random grouping, participants under six different experiment manipulations (explained below) were led to six different rooms to complete the experimental tasks independently.

We employed a 3 (PFP intensity: high, moderate, low) × 2 (performance evaluation criteria: objective measures vs. subjective measures) cross-sectional design. The experimental task employed was based on an anagram-solving exercise developed by Belogolovsky, Bamberger, Alterman, and Wagner (2016) and also used in Bamberger and Belogolovsky's (2017) experimental study on PFP. We combined this task with an essay-writing task to generate a new creative writing task to enable manipulations for both objective and subjective performance (see "Manipulations and Measures"). This 40-minute task involved anagram solving and essay writing. In each team, each participant was presented with a unique version of a 30-item anagram. Following Bamberger and Belogolovsky (2017), we allowed participants to seek help from team members in solving their own anagrams and to borrow other individuals' resolved words in writing

FIGURE 5
The Interactive Effects of PFP with Helping–Performance Expectancy on Employee Helping Behavior in Study 2



Note: HPE indicates helping–performance expectancy. PFP intensity ranged from 1 (0–5%) to 6 (70–85%) in Study 2.

their essays. To increase external validity (e.g., helping others might temporally conflict with performing individual tasks), the participants were forbidden from making alliances (e.g., exchanging anagrams with each other and solving all team members' anagrams together). Instead, we asked the participants to write on a "helping card" to lend help to other individuals either to solve the other individuals' anagrams or lend their own solved words to other individuals. Each participant had only 10 helping cards, and each card could be used to deliver only one word.

Manipulations and Measures

PFP. We manipulated three conditions of PFP with varying levels of incentive intensity (20%, 50%, and 80%). To maximize the experiment's external validity, the average pay level for a 40-minute task (plus a 20-minute survey) among four participants in a team was set to 50 Chinese yuan (approximately 7 U.S. dollars). In the high PFP condition, the participants were informed that they would receive 10 yuan in fixed pay as a reward for their participation. In addition, based on the outcome of individual performance rated by the leader, they would additionally receive 64, 48, 32, or 16 yuan (for performance grades A, B, C, and D, respectively) as a reward for their individual performance. In the moderate PFP condition, the fixed pay was 25 yuan, and the additional performance-based pay was 40, 30, 20, and 10 yuan (for performance grades A, B, C, and D, respectively). Finally, in the low PFP condition, the fixed pay was 40 yuan, and the additional performance-based pay was 16, 12, 8, and 4 yuan (for performance grades A, B, C, and D, respectively).

Performance subjectivity. Following Belogolovsky and Bamberger (2014), we manipulated performance subjectivity by randomly dividing the participants into two performance evaluation conditions: objective measures versus subjective measures. The participants in the objective condition were told that the rating of the task performance would be calculated objectively, entirely contingent on the number of solved words used in their completed creative essays. The number of solved anagrams and the length of the essay, which are also objective indicators, would be considered when two or more participants used the same number of solved words in writing their essays. In contrast, the

participants in the subjective condition were told that the rating of the task performance would be assessed subjectively based on an array of quantitative (e.g., number of used words solved from the anagrams and length of essay) and qualitative (e.g., logical consistency and creativity) indicators of the essay and the leader's observation of the individual's performance. Under both conditions, individual performance received one of four performance grades (A–D), and each performance grade should have been given to no more than one team member.

After completing the task, the participants were asked to complete a questionnaire that included measures of helping–performance expectancy, team members' helping behavior, and manipulation checks. All of the following scales were measured using a seven-point Likert scale (1 = "strongly disagree" to 7 = "strongly agree").

Helping–performance expectancy. Helping–performance expectancy was measured using the same 2-item scale employed in Study 2, and the reliability scores of these two items was $\alpha = .89$ and $\omega = .89$ in this study.

Helping behavior. We used two approaches to measure the participants' helping behaviors during the task. First, we employed a social network approach and asked participants to rate the extent to which each of their teammates lent help to them during the task. Considering the complexity in answering a round-robin questionnaire and the fitness of measures with our experimental context, we used a three-item scale of receiving help developed by Uy, Lin, and Ilies (2017). A sample item was "Team member *X* went out of their way to help me in the task" ($\alpha = .74$; $\omega = .76$). We then calculated each participant's helping behavior by averaging all other team members' scores on their perceived help received from the focal person for these three items. Second, we counted the number of helping cards each participant used in helping other individuals solve their anagrams and lending their own solved words to other individuals during the task as an objective measure of helping behavior. Subjectively measured helping was highly correlated with objectively measured helping in this experiment ($r = .50$, $p < .001$).

Manipulation check. The two manipulations in this experiment were both objective HR policies, and we utilized two methods to guarantee that the participants accurately understood these policies. First, before the task started, we provided each participant with a job description specifying the performance evaluation criteria, rating method, and reward policy (e.g., PFP intensity). Second, we asked the team leaders to explain these HR practices to each team member directly and to ensure that the participants' understandings were correct.

The leaders asked every team member to answer the following question to check the validity of the PFP manipulation: "On average, what is the proportion of the performance-based bonus in the total monetary reward for this task among all team members?" The participants answered correctly as follows: 80% under the high PFP condition, 50% under the moderate PFP condition, and 20% under the low PFP condition.

Performance subjectivity was measured using the same 3-item scale employed in Study 2. The reliability of these three items was .72 (α) and .75 (ω) in this study. Participants under the performance subjectivity condition ($M = 4.20$, $SD = .98$) gave higher scores on these measures than did participants under the performance objectivity condition ($M = 3.07$, $SD = 1.04$, $t(136) = 6.57$, $p < .001$). These results together indicate that all manipulations were effective.

Results

Given that PFP was manipulated as an ordinal multicategorical variable that consisted of three different experimental conditions, we utilized Hayes and Montoya's (2017) recommended regression analysis to test moderation effects for multicategorical independent variables. Specifically, as per Darlington and Hayes (2017), we adopted two different approaches (i.e., Helmert coding and indicator coding) to create two sets of dummies variables for the three PFP conditions manipulated in this study. Regression coefficients generated via Helmert coding were able to test the difference in helping behavior between the low PFP condition and the unweighted means for the moderate and high PFP conditions (coefficient of Helmert-*D*1) and the difference in helping behavior between the means of the moderate and high PFP conditions (coefficient of Helmert-*D*2). In addition, regression coefficients generated by indicator coding were able to test the difference in helping between the means of the low and moderate PFP conditions (coefficient of indicator-*D*1) and the difference between the means of the low and high PFP conditions (coefficient of indicator-*D*2). By employing these two coding approaches (see Schabram, Robinson, & Cruz, 2018), our analyses provided more nuanced findings regarding the differences in helping behavior across various comparison conditions. Table 6 shows the descriptive statistics and correlations for all variables in Study 3.

TABLE 6
Descriptive Statistics and Correlations of Variables in Study 3

Variables	Mean	*SD*	Correlations							
			1	2	3	4	5	6	7	8
1 Objective helping behavior	4.96	3.12	—							
2 Subjective helping behavior	5.03	.99	.50***	(.74)						
3 Performance subjectivity	.49	.50	–.04	.16†	—					
4 HPE	5.43	1.02	.20*	.10	.18*	(.89)				
5 Helmert-*D*1	–.00	.47	.08	.02	–.03	–.16	—			
6 Helmert-*D*2	.01	.41	.14	–.08	.03	–.08	.02	—		
7 Indicator-*D*1	.32	.47	–.08	.08	–.04	–.01	.49***	–.86***	—	
8 Indicator-*D*2	.34	.48	.16	–.06	.01	–.15	.51***	.87***	–.50***	—

Notes: $n = 140$. Dummy coding of multicategorical independent variable is as follows: Helmert-*D*1 = low PFP condition (–2/3), moderate PFP condition (1/3), high PFP condition (1/3); Helmert-D2 = low PFP condition (0), moderate PFP condition (–1/2), high PFP condition (1/2); Indicator-D1 = low PFP condition (0), moderate PFP condition (1), high PFP condition (0); Indicator-D2 = low PFP condition (0), moderate PFP condition (0), high PFP condition (1). HPE indicates helping–performance expectancy. Reliability coefficients are displayed in parentheses on the diagonal.

* $p < .05$
** $p < .01$
*** $p < .001$

Testing of Hypothesis 1 was conducted using the PROCESS macro for SPSS with 1,000 resamples (Hayes, 2012; Hayes & Montoya, 2017). Regression results are presented in Table 7. When using objective helping behavior as the dependent variable, the results suggested that in the condition of high performance subjectivity (PS = 1), objective helping behavior among participants in the high PFP condition ($M = 6.38$) was higher than that among the participants in the low ($M = 4.38$: indicator-*D*2 = 2.00, $SE = .88$, $p < .05$, 95% CI = [.25, 3.75]) and moderate ($M = 3.57$: Helmert-*D*2 = 2.80, $SE = .91$, $p < .01$, 95% CI = [1.00, 4.61]) PFP conditions. The difference in objective helping between the moderate and low PFP conditions, however, was not significant (indicator-*D*1 = –.80, $SE = .91$, n.s., 95% CI = [–2.61, 1.00]). In the condition of objective performance measurement (PS = 0), the results suggested that the differences in objective helping between the high ($M = 4.96$) and low ($M = 4.83$) PFP conditions (indicator-*D*2 = .13, $SE = .89$, n.s., 95% CI = [–1.63, 1.90]), the high and moderate ($M = 5.50$) PFP conditions (Helmert-*D*2 = –.54, $SE = .88$, n.s., 95% CI = [–2.29, 1.20]), and the moderate and low PFP conditions (indicator-*D*1 = .67, $SE = .89$, n.s., 95% CI = [–1.09, 2.44]) were all insignificant. These differences were contrasted and are graphed in Figure 6. Thus, Hypothesis 1 is again supported. When using subjective helping behavior as the dependent variable, however, these differences were no longer significant (see Table 7 for detailed information on the significance level of the regression coefficients). Altogether, these results suggest that performance subjectivity moderated the effect of PFP on individuals' objective helping behaviors but not subjective helping behaviors.

Hypothesis 2 proposed that performance subjectivity has a positive effect on helping–performance expectancy. We conducted an analysis of variance (ANOVA) to test this effect. The results suggested that individuals perceived helping–performance expectancy to be significantly higher (F (1, 138) = 4.68, $p < .05$, $\eta^2 = .03$) under the performance subjectivity condition ($M = 5.62$) than the performance objectivity condition ($M = 5.25$). Thus, Hypothesis 2 is supported.

Hypothesis 3, which suggested that PFP interacted with helping–performance expectancy to affect employee helping behavior, was tested using the PROCESS macro for SPSS with 1,000 resamples (Hayes, 2012; Hayes & Montoya, 2017). The results shown in Table 7 indicate that when individuals perceived that helping–performance expectancy was high (helping–performance expectancy = 6.45), objective helping behavior among participants under the high PFP condition ($M = 6.89$) was marginally higher (indicator-*D*2 = 1.58, $SE = .87$, $p < .10$, 95% CI = [–.13, 3.30]) than that among participants under the low PFP condition ($M = 5.31$) and significantly higher (Helmert-*D*2 = 2.25, $SE = .93$, $p < .05$, 95% CI = [.41, 4.10]) than that among participants under the moderate PFP condition ($M = 4.64$); however, the

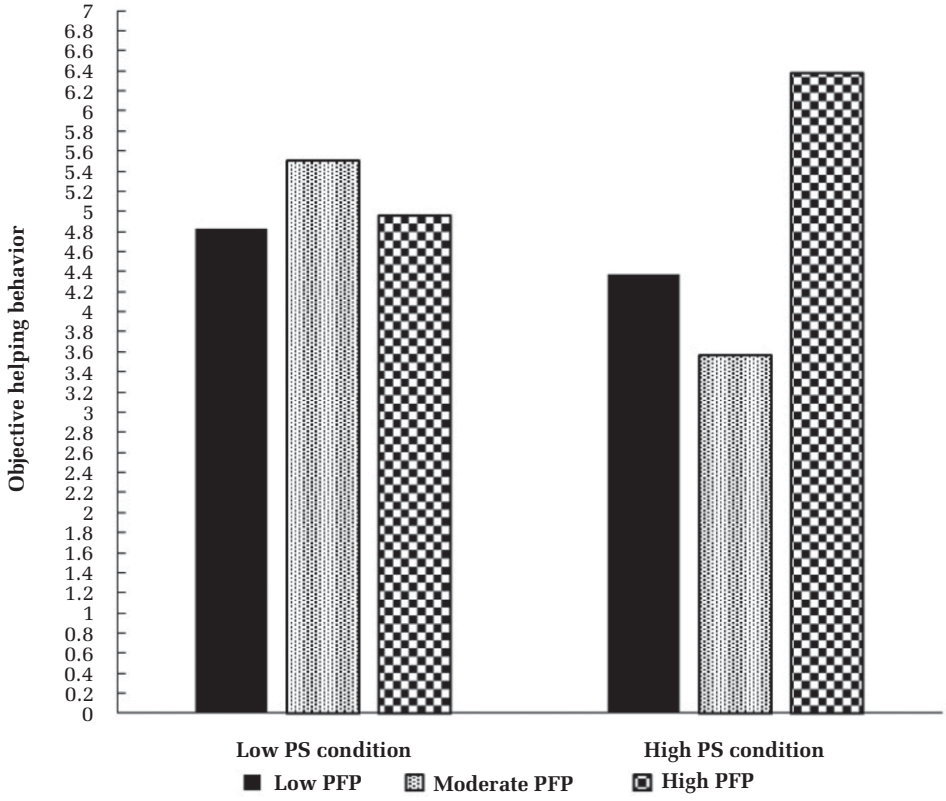
TABLE 7
Process Macro Analysis Results in Study 3

Variables	Objective Helping Behavior				Subjective Helping Behavior			
	High PS	Low PS	High HPE	Low HPE	High PS	Low PS	High HPE	Low HPE
Helmert-*D*1	.59 (.77) [–.93, 2.13]	.40 (.78) [–1.13, 1.94]	.46 (.75) [–1.02, 1.93]	1.24 (.86) [–.46, 2.95]	.30 (.25) [–.19, .78]	–.19 (.25) [–.68, .30]	.11 (.24) [–.36, .58]	.08 (.27) [–.47, .62]
Helmert-*D*2	2.80** (.91) [1.00, 4.61]	–.54 (.88) [–2.29, 1.20]	2.25* (.93) [.41, 4.10]	.29 (.85) [–1.39, 1.98]	.14 (.29) [–.44, .71]	–.55† (.28) [–1.10, .01]	.53† (.30) [–.06, 1.11]	–.77** (.27) [–1.30, –.23]
Indicator-*D*1	–.80 (.91) [–2.61, 1.00]	.67 (.89) [–1.09, 2.44]	–.67 (.89) [–2.44, 1.10]	1.10 (1.00) [–.88, 3.07]	.23 (.29) [–.35, .80]	.08 (.28) [–.48, .64]	–.15 (.28) [–.71, .41]	.46 (.32) [–.17, 1.09]
Indicator-*D*2	2.00* (.88) [.25, 3.75]	.13 (.89)[–1.63, 1.90]	1.58† (.87) [–.13, 3.30]	1.39 (.92) [–.43, 3.21]	.37 (.28) [–.19, .92]	–.46 (.28) [–1.02, .10]	.37 (.28) [–.17, .92]	–.31 (.29) [–.89, .27]

Notes: $n = 140$. Bootstrapping sample size = 1,000. PS = performance subjectivity. HPE = helping–performance expectancy. "High PS" reflects high PS condition (PS = 1). "Low PS" reflects low PS condition (PS = 0). "High HPE" reflects high HPE condition (HPE = mean + 1 $SD = 6.45$). "Low HPE" reflects low HPE condition (HPE = mean – 1 $SD = 4.40$). All coefficients displayed are unstandardized. Standard errors are in parentheses; 95% confidence intervals are in square brackets. As another test of the differences between groups, we reanalyzed the data using Mplus (verison 8.4; Muthén & Muthén, 2019), calculating the bias-corrected bootstrapped intervals for both the Helmert coding and the indicator coding dummy variables. We used the bias-corrected version of the bootstrapping method as recommended by MacKinnon, Lockwood, and Williams (2004), with 10,000 resamplings to ensure a high precision in the confidence interval. This robust analysis generated similar empirical results, which are available from the corresponding author.

$^{\dagger}p \leq .10$
$^{*}p \leq .05$
$^{**}p \leq .01$
$^{***}p \leq .001$

FIGURE 6
The Interactive Effects of PFP with Performance Subjectivity on Participant Objective Helping Behavior in Study 3



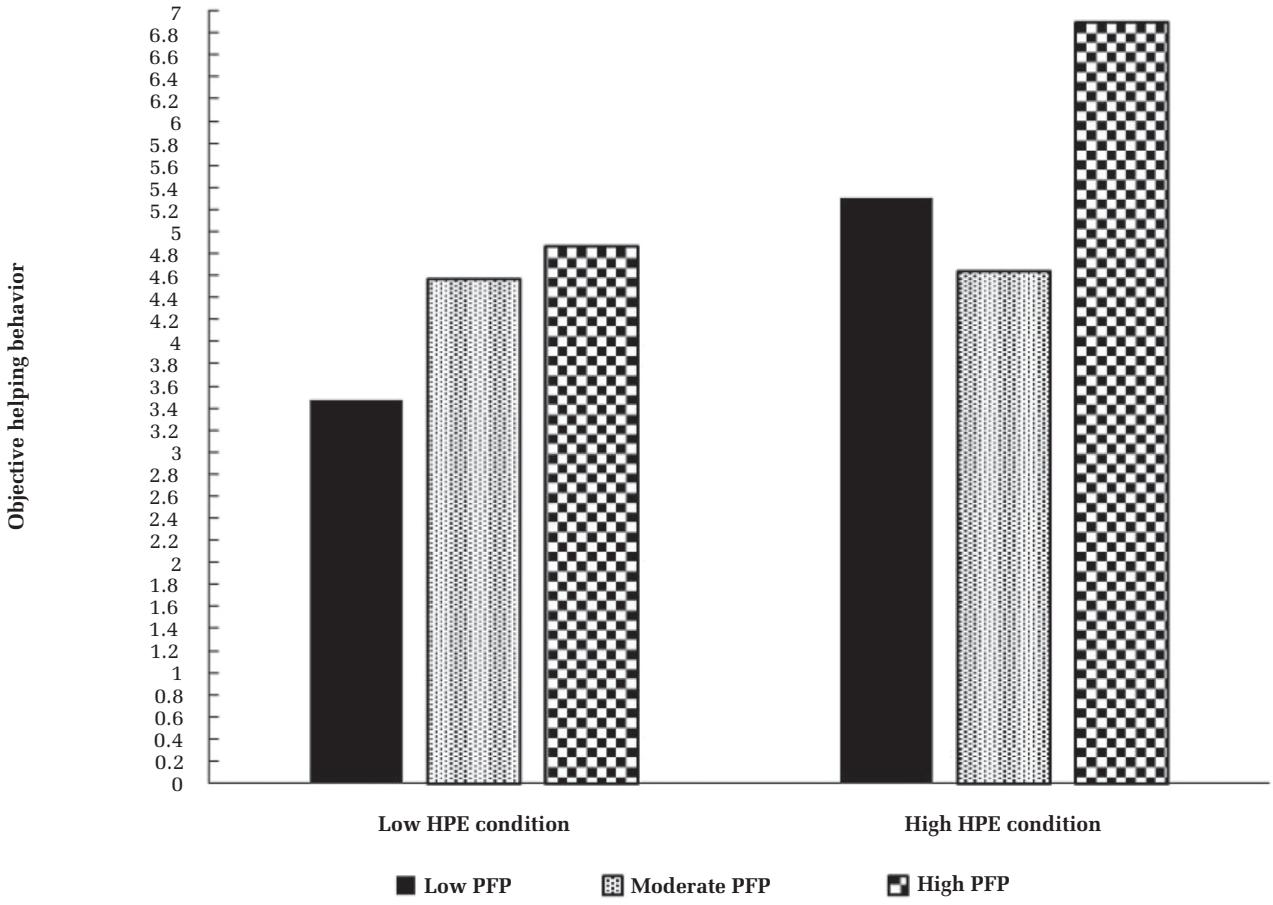
Note: PS = performance subjectivity.

difference in objective helping between the moderate and low PFP conditions was not significant (indicator-$D1 = -.67$, $SE = .89$, n.s., 95% CI = [–2.44, 1.10]). When individuals perceived that helping–performance expectancy was low (helping–performance expectancy = 4.40), the differences in objective helping between the high ($M = 4.86$) and low ($M = 3.47$) PFP conditions (indicator-$D2 = 1.39$, $SE = .92$, n.s., 95% CI = [–.43, 3.21]), the high and moderate ($M = 4.56$) PFP conditions (Helmert-$D2 = .29$, $SE = .85$, n.s., 95% CI = [–1.39, 1.98]), and the moderate and low PFP conditions (indicator-$D1 = 1.10$, $SE = 1.00$, n.s., 95% CI = [–.88, 3.07]) were all insignificant.

When using subjective helping behavior as the dependent variable, however, the differences across PFP conditions were no longer significant when the individual perceived helping–performance expectancy was high (see Table 7 for detailed information on the significance level of the regression coefficients). When individuals perceived that helping–performance expectancy was low, none of these differences were significant, except for the difference in subjective helping behavior between the high ($M = 4.66$) and moderate ($M = 5.43$) PFP conditions, indicating that PFP has a significant negative effect on subjective helping behavior (Helmert-$D2 = -.77$, $SE = .27$, $p < .01$, 95% CI = [–1.30, –.23]). Taken together, these results suggest that Hypothesis 3 is supported when using objective helping behavior as the dependent variable.[4]

[4] For robustness, we considered the review team's suggestion and reanalyzed our data by controlling each participant's performance grade as rated by the leader. While conducting analyses with and without control variables typically does not change the findings when a true underlying relationship is present, the practice does help rule out *p*-hacking as a cause for the findings (Sturman et al., in press). The empirical results, which are available from the corresponding author upon request, suggested that our findings remained unaffected.

FIGURE 7
The Interactive Effects of PFP with Helping–Performance Expectancy on Participant Objective Helping Behavior in Study 3



Note: HPE = helping–performance expectancy.

Discussion

Through a laboratory experiment, Study 3 helps to establish causality regarding the positive effect of performance subjectivity on perceived helping–performance expectancy and the interactive effects between PFP and performance subjectivity and between PFP and helping–performance expectancy on individual objective helping behavior. Moreover, the results suggest that the positive effect of PFP on helping behavior under the condition of high performance subjectivity is more likely to be significant at moderate to high levels of PFP intensity. Altogether, these findings suggest that the effect of PFP on individual helping behavior depends on both the degree of PFP intensity and how helping behavior is measured.

Despite the above causal effects demonstrated in Study 3, these experimental findings should be interpreted with caution due to the limitations of the experimental design. For instance, to effectively manipulate objective and subjective measures, additional performance requirements (e.g., logical consistency and creativity of the essay) were added under the subjective evaluation condition, leading to the possibility that the moderating effect of performance subjectivity was confounded by the influences of additional performance requirements. Although these additional performance requirements are theoretically unrelated to individual helping behavior, future research is encouraged to design different experimental tasks and different manipulations of performance measurement (subjectivity vs. objectivity) to replicate our findings. In addition, our measure of objective helping behavior may be biased by the issue of truncation—that is, the fact that each participant had a maximum of 10

helping cards. Nonetheless, given that the average number of helping card the participants used was only half of the total number of cards ($M = 4.96$; $SD = 3.12$), we are confident regarding the validity. Considering the pros and cons of each study, we discuss the general theoretical and practical implications of all three studies' findings below.

GENERAL DISCUSSION

The main purpose of our research was to challenge the conventional wisdom that PFP motivates employees' in-role task performance at the expense of undermining their motivations to engage in extra-role citizenship behaviors (e.g., coworker helping). Drawing upon expectancy theory, we highlighted the role of performance measurement and propose a motivation model that predicts the positive effect of PFP on employees' helping behaviors when employee performance is subjectively evaluated. The empirical results of the two field studies and an experimental study provided convergent support for the interactive effects between PFP and performance subjectivity on individual helping behavior. The results further indicated that the individual perceived helping–performance expectancy was the underlying psychological mechanism that accounted for why a pay-for-subjectively-evaluated-performance system can increase coworker helping. We next discuss the theoretical and practical implications of these findings.

Theoretical Implications

The first theoretical contribution of our research is that we demonstrate a positive effect of PFP on employee helping behavior in a context where individual performance is subjectively (as opposed to objectively) measured, evaluated, and rewarded. This finding provides nuanced insights into the inconclusive empirical results regarding the effect of PFP on individual helping behavior, as documented in the extant literature (Bamberger & Belogolovsky, 2017; Deckop et al., 1999, 2004; Du & Choi, 2010). Specifically, in a context in which individual performance is subjectively measured, the utilization of a PFP system may not cause conflict between employees' engaging in in-role tasks and engaging in helping behaviors; instead, PFP increases helping among colleagues (Deckop et al., 2004; Du & Choi, 2010). This conditional positive effect of PFP on helping behavior also enriches current understanding of the relational consequences of a PFP system in organizations (Belogolovsky & Bamberger, 2014; Nyberg et al., 2016).

Relatedly, we address past calls to pay more attention to the pivotal role of work context in advancing the theoretical development and empirical findings regarding PFP's effectiveness (Gupta & Shaw, 2014; Nyberg et al., 2016). Specifically, departing from prior studies that focused exclusively on the "pay" component of PFP (Gerhart et al., 2009; Rynes et al., 2005), we focus on the "performance" component of PFP and highlight the crucial context of performance appraisal (i.e., performance measurement) in shaping PFP's effects on individual work behaviors. The extant empirical evidence in the PFP literature "comes primarily from studies that use results-based (objective) measures of performance and from settings where work is simple, easy to measure, and not interdependent" (Gerhart & Fang, 2014: 42), ignoring the fact that employees' work performance and rewards may be highly dependent on their leaders' subjective evaluations. The economic approach in the PFP literature has increasingly emphasized the use of subjective measures in designing a more effective PFP system by preventing employees from "gaming" the objective performance indicators (Baker et al., 1994; Gibbs et al., 2004; Levin, 2003). Our research findings complement these economic studies by directly demonstrating the advantages (i.e., increased coworker helping) of using subjective measures in a PFP system, highlighting the importance of scrutinizing employees' enactment of PFP's signaling information in the workplace (Belogolovsky & Bamberger, 2014; Maltarich et al., 2017).

Notably, our findings do not suggest that subjective performance evaluations are superior to objective performance evaluations. In fact, the use of subjective performance evaluations may give rise to fairness problems and additional management costs due to managers' intentional and unintentional (e.g., heuristic) biases in evaluating employee performance, as reported in prior studies (Gerhart et al., 2009). Future research can contribute to the performance management literature by exploring the boundary conditions of the effectiveness of different performance measures to understand when subjective or objective measures appear to be a better choice. Notably, although the results of Study 3 suggest that the beneficial outcomes associated with the use of subjective performance measures may outweigh their inferior results only when the PFP intensity is at moderate to high levels, the companies surveyed in Studies 1 and 2, on average, still

employed a low to moderate level of PFP intensity. Therefore, more field data from various occupational, industrial, and cultural samples are needed to replicate and extend our empirical findings.

In addition, our research provides nuanced insight that enhances our understanding of the psychological mechanism (i.e., helping–performance expectancy) that transmits the interactive effects between the PFP intensity and performance subjectivity on employee helping behavior. Our research responds to recent calls for explorations of employees' psychological reactions under different PFP systems (Maltarich et al., 2017) and integrating interdisciplinary (e.g., economic and psychological) perspectives in examining PFP's contextualized influences (Larkin et al., 2012; Maltarich et al., 2017; Nyberg et al., 2016). The psychological mechanism of helping–performance expectancy as demonstrated in our research has notable implications for future empirical explorations in the PFP literature. Prior PFP research drawing upon expectancy theory to examine PFP's incentive effect has exclusively focused on the motivational component of instrumentality—that is, the perceived strength of the linkage between individual performance and pay—and sought to identify the context in which PFP's incentive effect could be strengthened by enhancing employees' instrumentality perceptions (Maltarich et al., 2017; Nyberg et al., 2016). Our model extends this stream of research by emphasizing and expanding the motivational component of expectancy—that is, the belief that there is a linkage between work effort and performance (Vroom, 1964).

Relatedly, our conceptualization and examination of helping–performance expectancy contribute to the advancement of expectancy theory. Given that PFP intensity positively affects individual instrumentality perception, this finding provides strong support for the central tenet of expectancy theory—individual behavior is motivated by the interaction between, but not the mere summation of, one's instrumentality and expectancy perceptions (Vroom, 1964). Moreover, relying on the assumption that human behavior is subjectively rational, expectancy theory has been questioned for its ignorance of individual cognitive limits on rationality and being "fixated on the amount rather than the direction of effort" (Vroom, 1995: 20). We partially attribute the formation of helping–performance expectancy to the individual heuristic biases activated in the context of high performance subjectivity. Our integration of the heuristic processing literature (Kahneman, 2011) provides new insights with respect to applying expectancy theory to predict the direction, rather than only the amount, of individual behavior.

Finally, we suggest that expectancy perceptions are not restricted to in-role work effort alone; instead, these perceptions originate from various behavioral domains (e.g., helping behavior) that are assumed to boost one's performance. By explicitly or implicitly incorporating extra-role behavioral domains into the performance appraisal system, organizations employing related PFP systems can motivate desired behaviors by enhancing domain-specific expectancy perceptions. Given that a large stream of PFP research has examined PFP effectiveness through an intrinsic–extrinsic motivation lens (Gerhart & Fang, 2015), our research redirects scholarly attention to underscore the positive implications of employees' domain-specific motivations related to various workplace behaviors.

Practical Implications

Beyond our theoretical implications, two key practical implications should be highlighted. First, although prior research has suggested that subjective measures suffer from "idiosyncratic rater effects" (Buckingham & Goodall, 2015), undermine the incentive effect of a PFP system (Belogolovsky & Bamberger, 2014; Murphy & Cleveland, 1995), and introduce additional financial and relational costs to the organization due to so-called "administrative problems" (Gerhart & Milkovich, 1992), our research found one possible benefit of using subjective measures in the PFP system—motivating employee helping behavior. By weighing the pros and cons, we suggest that a performance appraisal system that employs hybrid indicators, including both objective and subjective measures, is likely to motivate overall outcomes. For example, a balanced scorecard (Kaplan & Norton, 1992), if used appropriately, has better potential than the KPI technique, which is based purely on objective and financial measures, to promote employees' comprehensive work outcomes (Ittner, Larcker, & Meyer, 2003). Our results, however, should be interpreted with caution in practice; when launching a behavior-oriented performance appraisal method (e.g., behaviorally anchored rating scales [Bernardin & Smith, 1981]) among employees whose performance outcomes are not results-based or objectively measured, leaders should be systematically trained to avoid personal biases and evaluate employees' performances as objectively as possible.

Second, our motivational model of PFP and helping suggests that individual perceived helping–performance expectancy is the core psychological

impetus for helping behavior. Thus, organizations can motivate employees to perform above and beyond their prescribed in-role task obligations by activating their belief that "being a good soldier" is associated with desirable individual outcomes and monetary rewards. For example, organizations can incorporate extra-role behaviors, such as coworker helping, into their formal performance appraisal system to strengthen employee perceived helping–performance expectancy. In addition, organizations can design further incentive practices that directly reward employees' extra-role behaviors, such as a reward-for-voice plan, and bonuses for extraordinary extra-role contributions.

Limitations and Directions for Future Research

Despite these theoretical and practical contributions, several limitations of this research should be noted. First, although we operationalized PFP and performance subjectivity differently across the three studies in our research, the PFP system examined was short-term rather than long-term oriented (e.g., deferred compensation or stock option plans). Additionally, we considered coworker helping a within-team phenomenon in our research, and we did not consider team-level influences on employees' helping behaviors in our conceptual model. Our data were also unable to differentiate employees' within-team helping behaviors from their cross-team helping behaviors. Given that team leaders are typically the raters of employee performance, it is likely that our motivational model can be applied only to predicting employees' within-team helping behaviors. Notably, the motivational model of PFP and helping behavior in Study 3, which was the only study conducted within a team setting, was not supported when helping behavior was subjectively evaluated by teammates. This finding is intriguing because teammates, as recipients of the help, should have accurate information regarding all team members' helping behaviors. One explanation for these insignificant results is that the experiment participants may have been short-term oriented, paid less attention to relationship building and social interactions, and, hence, have been less likely to value and recognize the helping behavior of other individuals. Thus, compared with leaders, coworkers in certain settings provide a downward-biased evaluation of their colleagues' instrumentally motivated helping behaviors. We encourage future research to incorporate a real team design to unveil the role of teams (and teammates) on the relationship between PFP and helping.

Although we controlled for the influences of collective PFP and perceived task interdependence, future research is encouraged to examine the influence of more contextual factors at the team level, such as leadership. Leaders' direct modeling and enactment of an organization's HR practices can enhance the effectiveness of HR practices because leaders have immediate and frequent social interactions with employees and, thus, are regarded as representatives of the organization (Purcell & Hutchinson, 2007). Positive leadership behaviors, such as transformational behaviors (Weller, Süß, Evanschitzky, & von Wangenheim, 2019) and visionary behaviors (Ateş, Tarakci, Porck, van Knippenberg, & Groenen, 2018), have been found to strengthen employee consensus regarding HR practices, which, in turn, facilitates the execution and effectiveness of these HR practices. In contrast, frontline managers' leadership behaviors can also lead employees to hold negative attitudes toward HR practices when these managers are not strategically aligned with the CEO (Ateş et al., 2018). PFP efficacy is also likely affected by leadership because leaders can affect employee work motivation by enacting different leadership behaviors (Han et al., 2015). One central premise of our model is the effect by which performance subjectivity enhances individual helping–performance expectancy. Nonetheless, this premise may be challenged if the leader's implicit beliefs in evaluating employee performance are strictly against extra-role behaviors. We encourage future studies to advance our model by adding the role of leadership.

We also encourage future research to perform more nuanced examinations of the "performance" component of PFP, particularly the performance appraisal criteria and accuracy in designing objective and subjective indicators, when exploring PFP's influences on employees' workplace behaviors. For example, although our research demonstrates the advantage of using subjective performance measures in facilitating PFP's positive effect on employee helping behavior, it is unknown whether, why, and when the use of objective performance measures may cause other problems. Recent work has suggested that the legitimacy and accuracy of objective performance indicators play a vital role in impacting the effectiveness of the performance appraisal and reward system (Giarratana et al., 2018; Manthei & Sliwka, 2013). Thus, we urge future studies to examine how different dimensions of performance evaluation (e.g., performance basis, measurability, and indicator accuracy or legitimacy) work separately and jointly to shape PFP's effectiveness.

Another theoretical limitation of our research is that we did not consider alternative mediating mechanisms in our conceptual model, such as justice perceptions associated with the implementation of PFP in organizations (Park, Kim, & Sung, 2017; Trevor & Wazeter, 2006). Moreover, recent research has underscored the affect-based mechanism in explaining PFP's influences on individual behaviors (Belogolovsky & Bamberger, 2014; Maltarich et al., 2017). Thus, further research should integrate cognitive and affective mechanisms to explore employees' reactions to a PFP system. We also encourage future research to extend our conceptual model by considering more nuanced helping motives, such as prosocial, organizational concern and impression management motives (Rioux & Penner, 2001).

A final limitation of our research is the lack of a direct comparison between one's in-role task performance and extra-role helping behavior as two parallel behavioral outcomes in a PFP system. Our empirical findings may still be robust, however, because we ruled out the confounding influences of the participants' subjective, relative performance on helping behavior in Study 2 and Study 3. We urge future research to measure individuals' objective task performance and helping behavior simultaneously, and explore whether the use of a system that offers pay for subjectively evaluated performance decreases one's objective task performance (as claimed in prior research) while increasing one's helping behavior (as found in our research). Methodologically, future research could benefit from adopting a longitudinal perspective and tracking changes in the PFP intensity and helping behaviors over time. Future research might also consider conducting field experiments, randomly assigning organizations to implement varying levels of PFP intensity with both objective and subjective performance measures and tracking how these systems impact employees' outcomes.

CONCLUSION

Compensation scholars have recently called for nuanced explorations of PFP's impacts on individual extra-role behaviors. Our research contributes to this burgeoning stream of research by theorizing and testing a motivational model of PFP and helping behavior. By challenging the conventional view that PFP undermines employee extra-role behavior, our findings indicate that PFP has a conditional positive effect on individual helping behavior in the context of high performance subjectivity. Moreover, our findings emphasize the importance of examining the psychological mechanisms (i.e., helping–performance expectancy) that underlie employees' motivation to help in response to the PFP and performance appraisal systems that organizations enact.

REFERENCES

Aguinis, H. 2009. ***Performance management***. Upper Saddle River, NJ: Pearson Prentice Hall.

Aiken, L. S., & West, S. G. 1991. ***Multiple regression: Testing and interpreting interactions***. Newbury Park, CA: Sage.

Amabile, T. M. 1988. A model of creativity and innovation in organizations. ***Research in Organizational Behavior***, 10: 123–167.

Arvey, R. D., & Murphy, K. R. 1998. Performance evaluation in work settings. ***Annual Review of Psychology***, 49: 141–168.

Ateş, N. Y., Tarakci, M., Porck, J. P., van Knippenberg, D., & Groenen, P. J. F. 2018. The dark side of visionary leadership in strategy implementation: Strategic alignment, strategic consensus, and commitment. ***Journal of Management***, 46: 637–665.

Baker, G. 1992. Incentive contracts and performance measurement. ***Journal of Political Economy***, 100: 598–614.

Baker, G., Gibbons, R., & Murphy, K. J. 1994. Subjective performance measures in optimal incentive contracts. ***Quarterly Journal of Economics***, 109: 1125–1156.

Bamberger, P., & Belogolovsky, E. 2017. The dark side of transparency: How and when pay administration practices affect employee helping. ***Journal of Applied Psychology***, 102: 658–671.

Bamberger, P. A., & Levi, R. 2009. Team-based reward allocation structures and the helping behaviors of outcome-interdependent team members. ***Journal of Managerial Psychology***, 24: 300–327.

Belogolovsky, E., & Bamberger, P. A. 2014. Signaling in secret: Pay for performance and the incentive and sorting effects of pay secrecy. ***Academy of Management Journal***, 57: 1706–1733.

Belogolovsky, E., Bamberger, P., Alterman, V., & Wagner, D. T. 2016. Looking for assistance in the dark: Pay secrecy, expertise perceptions, and efficacious help seeking among members of newly formed virtual work groups. ***Journal of Business and Psychology***, 31: 459–477.

Bergeron, D. M. 2007. The potential paradox of organizational citizenship behavior: Good citizens at what cost? ***Academy of Management Review***, 32: 1078–1095.

Bernardin, H. J., & Smith, P. C. 1981. Clarification of some issues regarding the development and use of behaviorally anchored rating scales (BARS). ***Journal of Applied Psychology***, 66: 458–463.

Blau, G. 1994. Testing the effect of level and importance of pay referents on pay level satisfaction. ***Human Relations***, 47: 1251–1268.

Blau, P. M. 1964. ***Exchange and power in social life***. New York, NY: Wiley.

Bolino, M. C. 1999. Citizenship and impression management: Good soldiers or good actors? ***Academy of Management Review***, 24: 82–98.

Bolino, M. C., Turnley, W. H., Gilstrap, J. B., & Suazo, M. M. 2010. Citizenship under pressure: What's a "good soldier" to do? ***Journal of Organizational Behavior***, 31: 835–855.

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. B. 1995. On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. ***Personnel Psychology***, 48: 587–605.

Borman, W. C., & Motowidlo, S. J. 1993. Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), ***Personnel selection in organizations***: 77–98. San Francisco, CA: Jossey-Bass.

Brislin, R. W. 1986. The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), ***Field methods in cross-cultural research***: 137–164. Beverly Hills, CA: Sage.

Buckingham, M., & Goodall, A. 2015. Reinventing performance management. ***Harvard Business Review***, 93: 40–50.

Byron, K., & Khazanchi, S. 2012. Rewards and creative performance: A meta-analytic test of theoretically derived hypotheses. ***Psychological Bulletin***, 138: 809–830.

Cadsby, C. B., Fei, S., & Tapon, F. 2007. Sorting and incentive effects of pay for performance: An experimental investigation. ***Academy of Management Journal***, 50: 387–405.

Colquitt, J. A., & Simmering, M. J. 1998. Conscientiousness, goal orientation, and motivation to learn during the learning process: A longitudinal study. ***Journal of Applied Psychology***, 83: 654–665.

Darlington, R. B., & Hayes, A. F. 2017. ***Regression analysis and linear models: Concepts, applications, and implementation***. New York, NY: The Guilford Press.

Deci, E. L. 1975. ***Intrinsic motivation***. New York, NY: Plenum.

Deci, E. L., Koestner, R., & Ryan, R. M. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. ***Psychological Bulletin***, 25: 627–668.

Deci, E. L., & Ryan, R. M. 1985. ***Intrinsic motivation and self-determination in human behavior***. New York, NY: Plenum.

Deckop, J. R., Cirka, C. C., & Andersson, L. M. 2003. Doing unto others: The reciprocity of helping behavior in organizations. ***Journal of Business Ethics***, 47: 101–113.

Deckop, J. R., Mangel, R., & Cirka, C. 1999. Getting more than you pay for: Organizational citizenship behaviour and pay-for-performance plans. ***Academy of Management Journal***, 42: 420–428.

Deckop, J. R., Merriman, K. K., & Blau, G. 2004. Impact of variable risk preferences on the effectiveness of control by pay. ***Journal of Occupational and Organizational Psychology***, 77: 63–80.

De Dreu, C. K. W. 2006. Rational self-interest and other orientation in organizational behavior: A critical appraisal and extension of Meglino and Korsgaard (2004). ***Journal of Applied Psychology***, 92: 1625–1637.

Deming, W. E. 1986. ***Out of the crisis***. Cambridge, MA: Center for Advanced Engineering Study, MIT.

Du, J., & Choi, J. N. 2010. Pay for performance in emerging markets: Insights from China. ***Journal of International Business Studies***, 41: 671–689.

Eisenberger, R., & Cameron, J. 1996. Detrimental effects of reward: Reality or myth? ***American Psychologist***, 51: 1153–1166.

Eisenhardt, K. M. 1989. Agency theory: An assessment and review. ***Academy of Management Review***, 14: 57–74.

Flynn, F. J. 2006. How much is it worth to you? Subjective evaluations of help in organizations. ***Research in Organizational Behavior***, 27: 133–174.

Frey, B. S., & Osterloh, M. 2012. Stop tying pay to performance. ***Harvard Business Review***, 90: 51–52.

Fulmer, I. S., & Shaw, J. D. 2018. Person-based differences in pay reactions: A compensation-activation theory and integrative conceptual view. ***Journal of Applied Psychology***, 103: 939–958.

Garbers, Y., & Konradt, U. 2014. The effect of financial incentives on performance: A quantitative review of individual and team-based financial incentives. ***Journal of Occupational and Organizational Psychology***, 87: 102–137.

Gerhart, B. 2017. Incentives and pay for performance in the workplace. ***Advances in Motivation Science***, 4: 91–140.

Gerhart, B., & Fang, M. 2014. Pay for (individual) performance: Issues, claims, evidence and the role of sorting

effects. ***Human Resource Management Review***, 24: 41–52.

Gerhart, B., & Fang, M. 2015. Pay, intrinsic motivation, extrinsic motivation, performance, and creativity in the workplace: Revisiting long-held beliefs. ***Annual Review of Organizational Psychology and Organizational Behavior***, 2: 489–521.

Gerhart, B., & Milkovich, G. T. 1992. Employee compensation: Research and practice. In M. D. Dunnette & L. M. Hough (Eds.), ***Handbook of industrial & organizational psychology*** (2nd ed.): 481–569. Palo Alto, CA: Consulting Psychologists Press.

Gerhart, B., & Newman, J. M. 2020. ***Compensation*** (13th ed.). New York, NY: McGraw Hill.

Gerhart, B., & Rynes, S. 2003. ***Compensation: Theory, evidence, and strategic implications***. Thousand Oaks, CA: Sage.

Gerhart, B., Rynes, S. L., & Fulmer, I. S. 2009. Pay and performance: Individuals, groups, and executives. ***Academy of Management Annals***, 3: 251–315.

Giarratana, M. S., Mariani, M., & Weller, I. 2018. Rewards for patents and inventor behaviors in industrial research. ***Academy of Management Journal***, 61: 264–292.

Gibbs, M., Merchant, K. A., Van der Stede, W. A., & Vargus, M. E. 2004. Determinants and effects of subjectivity in incentives. ***Accounting Review***, 79: 409–436.

Gupta, A. K., & Govindarajan, V. 1984. Business unit strategy, managerial characteristics, and business unit effectiveness at strategy implementation. ***Academy of Management Journal***, 27: 25–41.

Gupta, N., & Shaw, J. D. 2014. Employee compensation: The neglected area of HRM research. ***Human Resource Management Review***, 24: 1–4.

Han, J. H., Bartol, K. M., & Kim, S. 2015. Tightening up the performance-pay linkage: Roles of contingent reward leadership and profit-sharing in the cross-level influence of individual pay-for-performance. ***Journal of Applied Psychology***, 100: 417–430.

Hayes, A. F. 2012. PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [Computer software]. Retrieved from http://www.afhayes.com/public/process2012.pdf

Hayes, A. F., & Montoya, A. K. 2017. A tutorial on testing, visualizing, and probing an interaction involving a multicategorical variable in linear regression analysis. ***Communication Methods and Measures***, 11: 1–30.

Hofmann, D. A., & Gavin, M. B. 1998. Centering decisions in hierarchical linear models: Implications for research in organizations. ***Journal of Management***, 24: 623–641.

Hofmann, D. A., Griffin, M. A., & Gavin, M. B. 2000. The application of hierarchical linear modeling to organizational research. In K. J. Klein & S. W. J. Kozlowski (Eds.), ***Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions***: 467–511. San Francisco, CA: Jossey-Bass.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. ***Accounting Review***, 78: 725–758.

Jenkins, D. G., Jr., Mitra, A., Gupta, N., & Shaw, J. D. 1998. Are financial incentives related to performance? A meta-analytic review of empirical research. ***Journal of Applied Psychology***, 83: 777–787.

Jensen, M. C., & Meckling, W. H. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. ***Journal of Financial Economics***, 3: 305–360.

John, L. S. K., Loewenstein, G., & Rick, S. I. 2014. Cheating more for less: Upward social comparisons motivate the poorly compensated to cheat. ***Organizational Behavior and Human Decision Processes***, 123: 101–109.

Kahneman, D. 2011. ***Thinking fast and slow***. New York, NY: Farrar, Straus and Giroux.

Kamdar, D., McAllister, D. J., & Turban, D. B. 2006. "All in a day's work": How follower individual differences and justice perceptions predict OCB role definitions and behavior. ***Journal of Applied Psychology***, 91: 841–855.

Kaplan, R. S., & Norton, D. P. 1992. The balanced scorecard—Measures that drive performance. ***Harvard Business Review***, 70: 71–79.

Kohn, A., Stewart, G. B., & McAdams, J. 1993. Rethinking rewards. ***Harvard Business Review***, 71: 48–49.

Kouchaki, M., Smith-Crowe, K., Brief, A. P., & Sousa, C. 2013. Seeing green: Mere exposure to money triggers a business decision frame and unethical outcomes. ***Organizational Behavior and Human Decision Processes***, 121: 53–61.

Kovacic, M. P., Galic, Z., & Jerneic, Z. 2014. Social desirability scales as indicators of self-enhancement and impression management. ***Journal of Personality Assessment***, 96: 532–543.

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

Lawler, E. E., III. 1971. ***Pay and organizational effectiveness: A psychological view***. New York, NY: McGraw-Hill.

Lawler, E. E., III. 2000. ***Rewarding excellence: Pay strategies for the new economy***. San Francisco, CA: Jossey-Bass.

Lawler, E. E., III, & Suttle, J. L. 1973. Expectancy theory and job behavior. ***Organizational Behavior and Human Performance***, 9: 482–503.

Levin, J. 2003. Relational incentive contracts. ***American Economic Review***, 93: 835–857.

Li, N., Zhao, H. H., Walter, S. L., Zhang, X. A., & Yu, J. 2015. Achieving more with less: Extra milers' behavioral influences in teams. ***Journal of Applied Psychology***, 100: 1025–1039.

MacKenzie, S. B., Podsakoff, P. M., & Fetter, R. 1993. The impact of organizational citizenship behavior on evaluations of salesperson performance. ***Journal of Marketing***, 57: 70–80.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. 2004. Confidence limits for the indirect effect: Distribution of the product and resampling methods. ***Multivariate Behavioral Research***, 39: 99–128.

Maltarich, M. A., Nyberg, A. J., Reilly, G., Abdulsalam, D. D., & Martin, M. 2017. Pay-for-performance, sometimes: An interdisciplinary approach to integrating economic rationality with psychological emotion to predict individual performance. ***Academy of Management Journal***, 60: 2155–2174.

Manthei, K., & Sliwka, D. 2013. ***Multitasking and the benefits of objective performance measurement—Evidence from a field experiment***. Working Paper, University of Cologne. Retrieved from https://www.econstor.eu/handle/10419/79968.

Marchegiani, L., Reggiani, T., & Rizzolli, M. 2016. Loss averse agents and lenient supervisors in performance appraisal. ***Journal of Economic Behavior & Organization***, 131: 183–197.

McNeish, D. 2018. Thanks coefficient alpha, we'll take it from here. ***Psychological Methods***, 23: 412–433.

Milgrom, P., & Roberts, J. 1992. ***Economics, organization and management***. Englewood Cliffs, NJ: Prentice Hall.

Mossholder, K. W., Richardson, H. A., & Settoon, R. P. 2011. Human resource systems and helping in organizations: A relational perspective. ***Academy of Management Review***, 36: 33–52.

Muckler, F. A., & Seven, S. A. 1992. Selecting performance measures: "Objective" versus "subjective" measurement. ***Human Factors***, 34: 441–455.

Murphy, K. R., & Balzer, W. K. 1986. Systematic distortions in memory based behavior ratings and performance evaluations: Consequences for rating accuracy. ***Journal of Applied Psychology***, 71: 39–44.

Murphy, K. R., & Cleveland, J. 1995. ***Understanding performance appraisal: Social, organizational, and goal-based perspectives***. Thousand Oaks, CA: Sage Publications.

Murphy, K. R., & DeNisi, A. 2008. A model of the performance appraisal process. In A. Varma, P. Budhwar, & A. DeNisi (Eds.), ***Performance management systems around the globe***: 81–96. London, U.K.: Routledge.

Muthén, L. K., & Muthén, B. O. 2019. ***Mplus user's guide***: 1998–2019. Los Angeles, CA: Muthén & Muthén.

Nyberg, A. J., Maltarich, M. A., Abdulsalam, D. D., Essman, S. M., & Cragun, O. 2018. Collective pay for performance: A cross-disciplinary review and meta-analysis. ***Journal of Management***, 44: 2433–2472.

Nyberg, A. J., Pieper, J. R., & Trevor, C. O. 2016. Pay-for-performance's effect on future employee performance: Integrating psychological and economic principles toward a contingency perspective. ***Journal of Management***, 42: 1753–1783.

Organ, D. W. 1988. ***Organizational citizenship behavior: The good soldier syndrome***. Lexington, MA: Lexington Books.

Organ, D. W. 1997. Organizational citizenship behavior: It's construct clean-up time. ***Human Performance***, 10: 85–97.

Organ, D. W., Podsakoff, P. M., & MacKenzie, S. B. 2006. ***Organizational citizenship behavior: Its nature, antecedents, and consequences***. Thousand Oaks, CA: Sage.

Park, S., & Sturman, M. C. 2016. Evaluating form and functionality of pay-for-performance plans: The relative incentive and sorting effects of merit pay, bonuses, and long-term incentives. ***Human Resource Management***, 55: 697–719.

Park, T.-Y., Kim, S., & Sung, L.-K. 2017. Fair pay dispersion: A regulatory focus theory view. ***Organizational Behavior and Human Decision Processes***, 142: 1–11.

Pearce, J. L., & Gregersen, H. B. 1991. Task interdependence and extrarole behavior: A test of the mediating effects of felt responsibility. ***Journal of Applied Psychology***, 76: 838–844.

Perlow, L., & Weeks, J. 2002. Who's helping whom? Layers of culture and workplace behavior. ***Journal of Organizational Behavior***, 23: 345–361.

Pfeffer, J. 1998. ***The human equation: Building profits by putting people first***. Boston, MA: Harvard Business School Press.

Podsakoff, P. M., Ahearne, M., & MacKenzie, S. B. 1997. Organizational citizenship behavior and the quantity and quality of work group performance. ***Academy of Management Journal***, 82: 262–270.

Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. 2000. Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. ***Journal of Management***, 26: 513–563.

Podsakoff, N. P., Whiting, S. W., Podsakoff, P. M., & Blume, B. D. 2009. Individual- and organizational-level consequences of organizational citizenship behaviors: A meta-analysis. ***Journal of Applied Psychology***, 94: 122–141.

Preacher, K. J., Curran, P. J., & Bauer, D. J. 2006. Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. ***Journal of Educational and Behavioral Statistics***, 31: 437–448.

Purcell, J., & Hutchinson, S. 2007. Front-line managers as agents in the HRM-performance causal chain: Theory, analysis and evidence. ***Human Resource Management Journal***, 17: 3–20.

Rioux, S. M., & Penner, L. A. 2001. The causes of organizational citizenship behavior: A motivational analysis. ***Journal of Applied Psychology***, 86: 1306–1314.

Roberson, L., Galvin, B. M., & Charles, A. C. 2007. When group identities matter: Bias in performance appraisal. ***Academy of Management Annals***, 1: 617–650.

Rotundo, M., & Sackett, P. R. 2002. The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: A policy-capturing approach. ***Journal of Applied Psychology***, 87: 66–80.

Rousseau, D. M. 1995. ***Psychological contracts in organizations: Understanding written and unwritten agreements***. Thousand Oaks, CA: Sage.

Ryan, R. M., & Deci, E. L. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. ***American Psychologist***, 55: 68–78.

Rynes, S. L., Gerhart, B., & Parks, L. 2005. Personnel psychology: Performance evaluation and pay for performance. ***Annual Review of Psychology***, 56: 571–600.

Schabram, K., Robinson, S. L., & Cruz, K. S. 2018. Honor among thieves: The interaction of team and member deviance on trust in the team. ***Journal of Applied Psychology***, 103: 1057–1066.

Shaw, J. D. 2015. Pay dispersion, sorting, and organizational performance. ***Academy of Management Discoveries***, 1: 165–179.

Shaw, J. D., & Gupta, N. 2007. Pay system characteristics and quit patterns of good, average, and poor performers. ***Personnel Psychology***, 60: 903–928.

Shaw, J. D., & Gupta, N. 2015. Let the evidence speak again! Financial incentives are more effective than we thought. ***Human Resource Management Journal***, 25: 281–293.

Shaw, J. D., Gupta, N., & Delery, J. E. 2002. Pay dispersion and workforce performance: Moderating effects of incentives and interdependence. ***Strategic Management Journal***, 23: 491–512.

Sturman, M. C. 2003. Searching for the inverted *U*-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. ***Journal of Management***, 29: 609–640.

Sturman, M. C., Cheramie, R. A., & Cashen, L. H. 2005. The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. ***Journal of Applied Psychology***, 90: 269–283.

Sturman, M. C., Sturman, A. J., & Sturman, C. J. In press. Uncontrolled control variables: The extent that a researcher's degrees of freedom with control variables increases various types of statistical errors. ***Journal of Applied Psychology***.

Super, J., Li, P., Ishqaidef, G., & Guthrie, J. 2016. Group rewards, group composition and information sharing: A motivated information processing perspective. ***Organizational Behavior and Human Decision Processes***, 134: 31–44.

Trevor, C. O., Reilly, G., & Gerhart, B. 2012. Reconsidering pay dispersion's effect on the performance of interdependent work: Reconciling sorting and pay inequality. ***Academy of Management Journal***, 55: 585–610.

Trevor, C. O., & Wazeter, D. L. 2006. A contingent view of reactions to objective pay conditions: Interdependence among pay structure characteristics and pay relative to internal and external referents. ***Journal of Applied Psychology***, 91: 1260–1275.

Tversky, A., & Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. ***Science***, 185: 1124–1131.

Uy, M. A., Lin, K. J., & Ilies, R. 2017. Is it better to give or receive? The role of help in buffering the depleting effects of surface acting. ***Academy of Management Journal***, 60: 1442–1461.

Van der Vegt, G. S., Bunderson, J. S., & Oosterhof, A. 2006. Expertness diversity and interpersonal helping in teams: Why those who need the most help end up getting the least. ***Academy of Management Journal***, 49: 877–893.

Van der Vegt, G. S., Van de Vliert, E., & Oosterhof, A. 2003. Informational dissimilarity and OCB: The role of intrateam interdependence and team identification. ***Academy of Management Journal***, 46: 715–727.

Van Dyne, L., Kamdar, D., & Joireman, J. 2008. In-role perceptions buffer the negative impact of low LMX on helping and enhance the positive impact of high LMX on voice. ***Journal of Applied Psychology***, 93: 1195–1207.

Van Dyne, L., & LePine, J. A. 1998. Helping and voice extra-role behaviors: Evidence of construct and predictive validity. ***Academy of Management Journal***, 41: 108–119.

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

Vroom, V. 1995. ***Work and motivation*** (1st edition). San Francisco, CA: Jossey-Bass.

Weller, I., Süß, J., Evanschitzky, H., & von Wangenheim, F. 2019. Transformational leadership, high-performance work system consensus, and customer satisfaction. ***Journal of Management***, 46: 1469–1497.

Whiting, S. W., Podsakoff, P. M., & Pierce, J. R. 2008. Effects of task performance, helping, voice, and organizational loyalty on performance appraisal ratings. ***Journal of Applied Psychology***, 93: 125–139.

Wei He (whe@nju.edu.cn) is an associate professor of management at the School of Business, Nanjing University, China. He received his PhD in business administration from Huazhong University of Science and Technology, China. His research interests include compensation, leadership, employee organizational citizenship behavior, and workplace deviance.

Shao-Long Li (tli@whu.edu.cn) is an assistant professor at the Economics and Management School of Wuhan University, China. He received his PhD in business administration from Huazhong University of Science and Technology, China. His research interests include compensation, team leadership, and workplace deviance.

Jie (Jasmine) Feng (jie.feng@rutgers.edu) is an assistant professor of human resource management in the Rutgers School of Management and Labor Relations. She received her PhD from the Department of Management and Human Resources at the Wisconsin School of Business. Her primary research interests include employee turnover, compensation, and interfaces between HR and entrepreneurship.

Guanglei Zhang (zhangguanglei@whut.edu.cn) is an associate professor at the School of Management, Wuhan University of Technology, China. He received his PhD in business administration from Huazhong University of Science and Technology, China. His research interests include human resource management practices and dirty work.

Michael C. Sturman (michael.sturman@rutgers.edu) is a professor of human resource management in the Rutgers School of Management and Labor Relations. He received his PhD from Cornell University's School of Industrial and Labor Relations. His research interests include the prediction of individual job performance over time, compensation systems, and HR analytics and metrics.