An Experimental Examination of Incentive and Sorting Effects of Pay-for-Performance on Creative Performance

by

Ji Hyun Kim

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Business)

at the
UNIVERSITY OF WISCONSIN-MADISON
2022

Date of final oral examination: 04/08/2022

This dissertation is approved by the following members of the Final Oral Committee:
Barry Gerhart, Professor, Management and Human Resources
Charlie Trevor, Professor, Management and Human Resources
Jirs Meuris, Assistant Professor, Management and Human Resources
Jee-Seon Kim, Professor, Educational Psychology

Acknowledgements

I would like to thank everyone who has helped me over the past five-year-long journey to completing a Ph.D. degree. First and foremost, I would like to extend my sincere gratitude and appreciation to my advisor, Barry Gerhart. I have had the fortune of being mentored by Barry, who is a successful scholar, wonderful advisor, and amazing role model of mine. His expertise, inspiration, resource support, feedback, patience, and humor have been paramount to my progress as a scholar, and I could never expect more.

Additionally, I would like to thank the members of my dissertation committee, Charlie Trevor, Jirs Meuris, and Jee-Seon Kim, for generously sharing their insights and time to better my work. I have been lucky to work alongside my friends and colleagues who are talented future scholars at the Wisconsin School of Business. Their energy, positivity, and enthusiasm were the sources of motivation when I was wandering the Grainger Hall.

Finally, I would like to express heartfelt gratitude to my mom and dad, Hye Ran Cho and Young Joe Kim, for their unconditional love, encouragement, and support. They have passed down the greatest gift of all – perseverance, diligence, kindness, and ambition. I owe my achievement and life to them.

Table of Contents

List of Tables

List of Figures

Abstract

There is a longstanding debate about whether pay-for-performance (PFP) enhances or undermines creative performance. Traditional motivation and revised creativity theories suggest that PFP and intrinsic task interest combine additively to enhance creative performance, whereas cognitive evaluation theory (CET) and self-determination theory (SDT) posit an undermining effect of PFP on task interest and thus on intrinsic motivation and creative performance. To help resolve these conflicting predictions and provide a more comprehensive understanding of how and when PFP influences creative performance, the current study incorporated both incentive and sorting mechanisms of PFP, varying strengths of PFP, and task autonomy as a key moderator. A novel laboratory experiment was designed to capture key elements of workplace contexts, including in the design of the creative tasks, the choice of PFP strength based on benchmarking of U.S. companies' practices, and allowing participants to sort into different pay conditions, consistent with the opportunity for mobility in the labor market. Results showed that, through both incentive and sorting mechanisms, high PFP intensity more strongly enhanced creative performance than did low PFP intensity and both were superior to fixed pay. Importantly, task autonomy positively moderated the PFP-creative performance relationship such that creative performance under PFP increased much more under higher task autonomy. Finally, the difference in creative performance under PFP versus fixed pay was greater when participants were allowed to sort into their preferred pay conditions than when they worked only under randomly assigned pay conditions (typical in past research on PFP and creativity). Several theoretical and practical implications and future directions are discussed.

Introduction

Although most organizations use pay-for-performance (PFP) to encourage employee effort and performance and attract and retain desired types of employees (Gerhart & Newman, 2020), there have been competing arguments concerning the effects of PFP on intrinsic motivation and thus on creative performance specifically (e.g., Byron & Khazanchi, 2012; deCharms, 1968; Deci, Olafsen, & Ryan, 2017; Eisenberger, Rhoades, & Cameron, 1999; Gerhart & Fang, 2015; Hennessey & Amabile, 2010). Traditional motivation theories (e.g., reinforcement theory, Skinner, 1953; equity theory, Adams, 1963; expectancy theory, Vroom, 1964) suggest a positive effect of PFP on performance generally, including creative performance, because motivation stemming from intrinsic task interest and extrinsic incentives combine additively (Gupta & Shaw, 1998; Jenkins, Mitra, Gupta, & Shaw, 1998; Kim, Gerhart, & Fang, 2022). In contrast, cognitive evaluation theory (CET) and later self-determination theory (SDT) posit that PFP, typically measured as an individual incentive, is detrimental to intrinsic motivation as extrinsic motivators are typically viewed as controlling and undermining one's self-determination, self-competence, and autonomy (Deci, Koestner, & Ryan, 1999; Deci & Ryan, 1985; Ryan & Deci, 2017). Thus, SDT suggests that incentives have an adverse effect on performance, particularly when a task requires greater intrinsic motivation and creativity where the risk of losing intrinsic interest is high (Deci & Ryan, 1985; Ryan & Deci, 2000a, 2017).

Similarly, the traditional theoretical view on creativity, which owes a significant intellectual debt to CET and SDT, argues that extrinsic motivators harm intrinsic interest and thus creative performance because intrinsic motivation and extrinsic motivation function in the opposite direction (Amabile, 1983a, 1998; Hennessey, 1989; Lepper, Greene, & Nisbett, 1973). However, Amabile and colleagues later proposed a "radically modified" creativity theory, which

makes the quite different prediction that extrinsic rewards may "bolster" intrinsic motivation and creativity if they are implemented in an autonomy- and information-supportive manner and/or if initial intrinsic motivation is already high (Amabile & Pratt, 2016; Hennessey & Amabile, 2010). Eisenberger and colleagues also found that rewards enhance intrinsic motivation and creativity by increasing perceived self-determination (Eisenberger, Pierce, & Cameron, 1999; Eisenberger & Rhoades, 2001; Eisenberger et al., 1999). Even with these evolving perspectives, there continues to be a theoretical disagreement over whether PFP should be implemented in work requiring creativity (e.g., Deci et al., 2017; Howard, Gagné, & Bureau, 2017; Ryan & Deci, 2017). Additionally, practitioners in various outlets, such as best-selling books and TED talks, have also consistently argued that PFP undermines employee creativity (e.g., "pay-for-performance schemes…usually don't work and often do harm," Pink, 2011; "To dramatically increase creativity, slash your budget in half," Schaefer, 2018; "How money kills creativity," Meyer, 2019). Hence, scholars have called for research that explicitly and objectively examines the effect of PFP on creative performance (Baer, Oldham, & Cummings, 2003; Erat & Gneezy, 2016; Gerhart & Fang, 2015; Kim et al., 2022; Park, Eaglesham, Shaw, & Burton, in press).

This longstanding debate about the PFP-creative performance relationship is especially important because creativity is regarded as one of the critical determinants of organizational success as it is seen as crucial to organizations becoming more innovative, competitive, and responsive to changes (Mumford & Gustafson, 1988; Shalley, Zhou, & Oldham, 2004; Zhou & Hoever, 2014). This is more so as the economy continues to evolve from reliance on labor-intensive work to knowledge-intensive work (Alvesson, 2004; Anad, Gardner, & Morris, 2007). Moreover, given that organizations have more discretion in *how* they compensate employees (e.g., *whether* and *to what extent* PFP is used) than in *how much* they compensate (Gerhart &

Milkovich, 1990), this may suggest, in turn, that organizations use this discretion in deciding how they pay to achieve higher employee creativity. However, the effects of PFP are less clear and indeed regarded by some theories to be especially harmful to creative performance. This is, as noted, because under traditional creativity theory "creativity is considered more dependent on intrinsic motivation, more cognitively demanding, more risky, and of less certain value than routine performance" (Byron & Khazanchi, 2012, p. 809). Consequently, the question of whether PFP undermines or enhances employee creativity is an important one, calling for better evidence-based guidance (Park et al., in press).

As such, the current study aims to help resolve the sharply competing theoretical predictions and differing practical implications regarding the effect of PFP on creative performance. To that end, I first examine the incentive effect of PFP by comparing creative performance under PFP versus fixed pay. Then, moving beyond the dichotomy of PFP and fixed pay, I investigate how (two) different strengths of PFP affect creative performance. Competing theoretical predictions of how PFP influences creative performance diverge more sharply with stronger PFP intensity (Kim et al., 2022). That is, according to the SDT logic, the detrimental effect of PFP on creative performance becomes more severe as PFP intensity increases. In contrast, based on traditional motivation (e.g., expectancy theory; Vroom, 1964) and revised creativity (Amabile & Pratt, 2016) theories, higher PFP intensity increases instrumentality more strongly and thus leads to superior performance, including creative performance, than lower PFP intensity does.

Importantly, it may be possible to reconcile these competing theoretical arguments, to a degree, by incorporating key theoretical moderators, such as task autonomy. Specifically, SDT and creativity theory posit that PFP may enhance creative performance when implemented in an

autonomy-supportive climate (Cerasoli, Nicklin, & Ford, 2014; Gagné & Deci, 2005; Hennessey & Amabile, 2010). However, no study, to my knowledge, has empirically investigated this important question using actual, objectively defined PFP schemes and creative performance. My study helps remedy that situation.

Furthermore, I test both the *incentive* and *sorting* effects of PFP and discuss how they function simultaneously to influence creative performance. Integrating the sorting effect with the incentive effect is particularly important in understanding the relationship between PFP, intrinsic motivation, and creative performance because sorting processes are likely to reduce poor matches between compensation schemes and workers, thus making any adverse effects of PFP less likely (Gerhart & Fang, 2015). An indication of the high degree of employee mobility in the labor market is that the annual total separation rate in the U.S. labor market in 2019 was 45%, with the majority being in the form of worker quits (28%; JOLTS: Job Openings and Labor Turnover Survey)[1]. This high employee mobility demonstrates a very substantial number of people changing their jobs to seek a better match (Davis & Haltiwanger, 1999; Jovanovic & Moffitt, 1990; Lazear & Spletzer, 2012). Unless such mobility is also allowed in experimental design, this important matching process will be missed. Nevertheless, the sorting effect of PFP has received much less attention than the incentive effect of PFP in the compensation, motivation, and creativity literatures. Although there are some key exceptions that incorporated both the incentive and sorting effects (Cadsby et al., 2007; Dohmen & Falk, 2011; Lazear, 2000; Trevor et al., 2012), none of them examined creativity as a performance measure.

[1] Total separations include quits (i.e., voluntary separations initiated by the employee), layoffs and discharges, and other separations. JOLTS reported that quits and layoffs and discharges account for 93% of total separations. After the break of COVID-19, the annual total separation and quit rates were 57% and 25%, respectively, in 2020 and 47% and 33%, respectively, in 2021.

Especially noteworthy is that the sorting effect has been completely overlooked in the traditional CET paradigm (e.g., Deci et al., 1999), in which the fundamental conceptual argument is that PFP harms intrinsic motivation and creativity. Any such harmful effects are made more likely in research using the CET paradigm where subjects are randomly assigned to PFP condition in the absence of opportunity for subjects to sort into their preferred pay condition, causing mismatches (Fang & Gerhart, 2012; Gerhart & Fang, 2015). My study seeks to determine whether any detrimental effects of PFP on creative performance are diminished or even eliminated when self-selection of PFP is allowed, just as workers change employers based on their preferences rather than being randomly assigned to and obligated to work (indefinitely) under PFP or fixed pay schemes, regardless of their preferences.

To test the proposed model, I designed a randomized, multi-stage laboratory experiment, which builds on previous research, but provides several important advantages. First, although field research has clear strengths (e.g., external validity; Shadish, Cook, & Campbell, 2002), studies conducted in field settings are often limited with respect to internal validity. With random assignment, stronger causal inference (internal validity) is often possible in a laboratory experiment than in a typical field setting (Shadish et al., 2002). Using a randomized laboratory experiment better enables isolation of the effects of independent and moderating variables of interest on outcome variables through superior control of confounding factors, both known and unknown. For instance, in his study of the effect of a new piece-rate scheme at an automobile windshield repair company, Lazear (2000) raised a concern that the observed productivity increase could be attributed to not only the firm's introduction of the new PFP scheme but also its transition to a new management team. My study's laboratory experiment design enables me to more validly identify clear cause-and-effect of the incentive and sorting mechanisms on creative

performance and rule out potential alternative explanations (Baer et al., 2003; Barnes, Hollenbeck, Jundt, DeRue, & Harmon, 2011).

Second, field research on PFP has used subjective, less quantifiable, measures of PFP or included less typical forms of rewards. Instead, they used individual-level perceptual measures of PFP, asking survey respondents to report how their employers compensated employee creativity (e.g., Baer et al., 2003; Malik, Butt, & Choi, 2015; Sue-Chan & Hempel, 2016; Yoon, Sung, & Choi, 2015). Yet, we know that the interrater reliability of such self-reports is typically quite low (Gerhart, 1999; Wright, Gardner, Moynihan, Park, Gerhart, & Delery, 2001), and even descriptions of human resource practices by expert informants have low reliability (Gerhart, Wright, McMahan, & Snell, 2000). Thus, such self-reports, rather than capturing true variance in PFP, may be significantly idiosyncratic to the individual.[2] In my experiment, in contrast, the PFP manipulation is quantifiable and clear, and I document subjective perceptions confirm that. Furthermore, some of the previous studies included questionnaires that measured broader forms of extrinsic rewards such as recognition, benefits, celebration dinner, and compliments (e.g., Malik et al., 2015; Sue-Chan & Hempel, 2016; Yoon et al., 2015) rather than PFP, per se. Another issue is that some previous studies used (perceptual) measures of PFP where pay was not contingent on creative performance specifically, but rather contingent on performance in general (e.g., Zhang, He, Long, & Zhang, 2022; Zhang, Long, & Zhang, 2015), which may further undermine internal validity to examine the linkage between PFP and creative performance. Finally, none of the previous (broadly defined) extrinsic rewards-creative

[2] As Gerhart and colleagues (1999, 2000) and Wright et al. (2001) pointed out, studies on human resource practices have focused on internal consistency reliability, while ignoring interrater reliability, which is typically a much larger source of measurement error. To my knowledge, no field studies on the extrinsic rewards-creativity relationship reported any form of interrater reliability. My study does.

performance studies conducted in the workplace incorporated the sorting effect of PFP on creative performance, even though, as I have documented, employee mobility is substantial. In sum, although these studies provide important building blocks in understanding the relationship between extrinsic rewards and creativity in the workplace, as Park et al. (in press) pointed out, "empirical research on incentive effects and creativity remain[s] limited" (p. 30).

Third, although laboratory experiments have the strengths described, the specific design of previous laboratory experiments on compensation, motivation, and creativity strongly limit the ability to draw implications for practice. For instance, as I describe in more detail later, key differences between my study design and that of previous laboratory experiments are the use of non-adult subjects in non-workplace settings (e.g., Boggiano & Ruble, 1979; Eisenberger & Rhoades, 2001), simple tasks less typical of the workplace such as generating titles for a short story or solving riddles (e.g., Eisenberger & Armeli, 1997; Kachelmeier, Reichert, & Williamson, 2008), and broadly defined extrinsic rewards that often do not include monetary compensation (e.g., Dodd & Ganster, 1996; Friedman, 2009).

Fourth, these previous laboratory experiments rely exclusively on a between-subjects design, which has the advantage of randomly assigning PFP conditions to subjects (e.g., Campbell, 1984; Gneezy & Rustichini, 2000) but does not allow subjects to sort into their preferred pay conditions, a departure from the reality of workplace settings and how the labor market operates. A study design combining between- and within-subjects factors is more desirable.

To overcome these limitations of prior research, I sought to design a laboratory experiment (see **Figure 1**) that aims to go beyond previous work by better incorporating key features of the workplace (Ilgen, 1986). I recruited adult participants and designed creative tasks

more in line with those found in the workplace. In addition, I provided a salient and more realistic level of PFP based on how U.S. companies actually pay (Rynes, Schwab, & Heneman III, 1983; WorldatWork, 2018). Lastly, I conducted a multiple round experiment, which included not only a between-subjects component, but importantly a within-subjects component as well, which allows investigating how creative performance changes under different pay conditions and allows subjects to sort and observe how people sort/make pay choices when they are able to choose PFP (and strength) as is the case in the actual labor market.

In summary, this study aims to contribute to the broad motivation, compensation, and creativity literatures by addressing a longstanding debate over the effect of PFP on creative performance. In so doing, first, I examine the incentive effect of PFP—whether PFP, in contrast to fixed pay, undermines or enhances creative performance. Then, I take a more refined approach by testing the varying strengths of PFP, as I predict that either the negative or positive effect of PFP will become greater as PFP intensity increases. Next, I examine task autonomy as a key moderator that may help resolve competing theoretical arguments. Specifically, based on SDT and creativity theory, I propose that even if there exists a negative PFP-creative performance main effect relationship, the relationship will be positive under high task autonomy. In contrast, if there already exists a positive PFP-creative performance relationship, as traditional motivation and revised creativity theories anticipate, it will become more positive as task autonomy increases. Finally, as noted, I incorporate the sorting effect of PFP to better mirror the workplace, where there is substantial worker mobility.

Theory and Hypotheses Development

Incentive Effects of PFP on Non-creative Performance

As summarized in **Figure 2**, CET (and later SDT), traditional motivation theories, and traditional and revised creativity theories take divergent perspectives on the effects of PFP on performance in different types of tasks: non-creative (noninteresting and interesting) and creative. It is crucial to distinguish the task types in understanding the arguments of these four influential theories because, according to some theories, different task types require different source of motivation to generate high performance. In addition, because the discussion of the PFP-performance relationship in the two types of non-creative tasks (noninteresting and interesting) provides the groundwork for addressing the relationship in creative tasks, I first review the theories and findings regarding the two types of non-creative tasks.

As seen most clearly under CET, performance in noninteresting and interesting tasks is driven by different types of motivation, extrinsic or intrinsic. Extrinsic motivation refers to "doing something because it leads to a separable outcome," such as financial incentives (Ryan & Deci, 2000, p. 55), whereas intrinsic motivation is "to engage in work primarily for its own sake, because the work itself is interesting, engaging, or in some way satisfying" (Amabile, Hill, Hennessey, & Tighe, 1994, p. 950). SDT and traditional motivation theories agree that PFP enhances performance in noninteresting tasks where extrinsic motivation is regarded to be a key driver of performance (Deci & Ryan, 1985; Latham, 2007). In support of this, previous meta-analyses (Jenkins et al., 1998; Kim et al., 2022; Weibel, Rost, & Osterloh, 2010) consistently found a positive PFP-performance relationship in noninteresting tasks, with performance being higher by .42 to .89 standard deviations using PFP (typically a piece-rate scheme, individual incentives).[3]

[3] δ (corrected Cohen's *d*) = .89 or $\hat{\rho}$ (mean estimated population correlation) = .35 in Jenkins et al. (1998); *d* (average sample size weighted Cohen's *d*) = .42 in Weibel et al. (2010); and δ = .52 in Kim et al. (2022)

In contrast, the theories disagree on the effect of PFP on performance in interesting tasks. As noted, traditional motivation theories such as expectancy theory (Vroom, 1964) predict that task interest (intrinsic motivation) and PFP (extrinsic motivation) combine additively to enhance performance. On the other hand, SDT suggests that PFP undermines performance in interesting tasks because extrinsic (e.g., financial incentives, PFP) and intrinsic motivators (e.g., inherent task interest) do *not* combine additively to increase total motivation (Deci et al., 1999; Deci & Ryan, 1985; Ryan & Deci, 2000b). Instead, SDT posits that extrinsic motivation is detrimental to intrinsic motivation, arguing that "while extrinsic rewards such as money can certainly motivate behavior, they appear to be doing so at the expense of intrinsic motivation" (Deci, 1972, p. 224-225). This is because PFP is believed to control one's behavior, evoking an external perceived locus of causality and eventually undermining self-determination and intrinsic motivation (deCharms, 1968; Deci et al., 2017; Ryan & Deci, 2017). Importantly, under the notion of SDT, this detrimental effect of PFP will be most critical in interesting tasks where initial intrinsic motivation at risk is already high, compared to noninteresting tasks where initial intrinsic motivation at risk is low.

To address the contradicting predictions on the effect of PFP on performance in interesting tasks, Jenkins et al. (1998) conducted a meta-analysis and found a positive effect of financial incentives on task performance in intrinsic (δ = .79 or $\hat{\rho}$ = .34) tasks. Later, however, Weibel et al.'s (2010) meta-analysis, although examining the same research question, found an effect size (d) of -.13 between PFP and task performance in intrinsically motivating tasks, reaching not only the opposite conclusion, but also being dramatically different in terms of effect size (.79 – (-.13) = .a 92 SD difference) from that of Jenkins et al. (1998), thus leaving the debate unresolved. Subsequently, Kim et al.'s (2022) more extensive meta-analysis (including studies

from Jenkins et al. (1998), from Weibel et al. (2010), as well as additional studies) found a positive relationship between financial incentives and performance in interesting (δ = .58) tasks. (Kim et al. (2022) also identified methodological reasons for the discrepancy between the two previous meta-analyses.)

Although the Kim et al.'s (2022) meta-analysis seems to resolve the controversy over the relationship of PFP with performance, they identified a key unanswered question—how PFP influences creative performance specifically. As their results showed substantially different effect sizes of PFP on performance quality versus quantity measures, they suggested that different types of performance may respond differently to PFP. As such, they called for future studies on the effect of PFP that examines creative performance particularly in settings as close workplace as possible, especially by use of (a) workplace typical PFP schemes (as a piece-rate scheme is actually not widely used in the workplace) (b) subjective evaluation of performance (the most common), and (c) an explicit, direct measure of varying strengths of incentive intensity (in contrast to the typical research design comparing one incentive strength to fixed pay). Lastly, they suggested future research (d) examines moderators of theoretical importance, especially autonomy and (e) considers the sorting effect of PFP, in addition to the incentive effect. The current study responds to this call by directly investigating the incentive effects of PFP intensity on creative performance and the moderating role of task autonomy on the relationship, in addition to the sorting effects of PFP as described in the following sections.

Incentive Effect of PFP on Creative Performance

It has been argued that because "creativity can benefit every function of an organization" (Amabile, 1998, p. 76) and contemporary organizations are increasingly relying on innovation and creativity to prosper (Anad et al., 2007; Shalley, Gilson, & Blum, 2009; Zhou & Hoever,

2014), it is crucial to understand how organizations can motivate employee creativity. By definition, creativity in a work setting indicates both *novelty* and *usefulness* of an idea (Amabile, 1998; Shalley et al., 2004; Zhou, 1998). Based on the componential model of creativity shown in **Figure 3,** unlike non-creative tasks, creative tasks require all of the three major components for superior performance: intrinsic motivation, domain-specific skills, and creativity-relevant processes and skills (Amabile, 1983a; Amabile & Pratt, 2016). Thus, it suggests that despite extant studies' focus on the influence of intrinsic motivation, other factors also contribute to creative performance, indicating that intrinsic motivation alone cannot predict creative performance (George, 2007; Grant & Berry, 2011). A second determinant of creativity is skills, expertise, and knowledge in a task domain or in multiple domains to generate the most novel and useful ideas. The third component is creativity-relevant processes and skills, which are necessary to utilize domain-specific skills in new, different ways. Furthermore, creative tasks demand heuristic approaches because paths to the solution are not clear or straightforward (Amabile, 1983b) and an appropriate "work environment" that can enhance these three components (Amabile & Pratt, 2016). Altogether, because creative tasks have such distinctive attributes, the arguments and findings regarding determinants of performance in non-creative tasks, including PFP, cannot be generalized to creative tasks.[4]

[4] As noted, the effect of PFP on performance in creative tasks has not been studied in workplace settings or in laboratory settings that effectively capture at least some key aspects (e.g., adult subjects, PFP, the nature of creative tasks) of workplace settings (Gerhart & Fang, 2015; Kim et al., 2022; Park et al., in press). Although theory and research on CET contributed early on to concerns in about the potential negative effects of PFP on creative performance, it is nevertheless important to keep in mind that the standard CET paradigm does not use performance of any type as a dependent variable. (Instead, the key dependent variable in the CET paradigm is how much time subjects spend on an interesting task during a free-time portion of the experiment.) Further, although the CET paradigm focuses on interesting tasks, a

The componential model of creativity is analogous to the ability-motivation-opportunity (AMO) framework, which suggests employee performance is a function of *ability*, *motivation*, and *opportunity* (Appelbaum, Bailey, Berg, & Kalleberg, 2000; Blumberg & Pringle, 1982; Gerhart, 2007). As such, I have added the AMO factors to **Figure 3**. Although research including the other components would be useful, my study begins with a focus on the motivation (M) component, which as we have seen, had until recently been equated with intrinsic motivation, omitting extrinsic motivation (PFP in particular). I aim to remedy that omission and to test competing theoretical predictions in this part of the model. Further, in the between-subjects part of the design, I randomly assign subjects to two equivalent creative tasks and two levels of task autonomy (a key theoretical moderator), which are important dimensions of the opportunity (O) dimension. Lastly, the ability (A) component is intentionally not controlled in the within-subjects part because I wish to allow people with different abilities and/or preferences to sort into PFP (or fixed pay) condition.

Early theoretical work on creativity was consistent with the argument from SDT. Amabile (1983a) suggested the Intrinsic Motivation Hypothesis of Creativity, stating that "the intrinsically motivated state is conducive to creativity, whereas the extrinsically motivated state is detrimental" (p. 15) because "extrinsic constraints … impair[ing] intrinsic motivation" (Amabile, 1983b, p. 366). Likewise, Amabile (1998) argued in her influential article that "people will be most creative when they feel motivated primarily by the interest, satisfaction, and challenge of the work itself—and not by external pressures" because employees may feel being

---

task, even if interesting, requires further properties (i.e., novelty and usefulness; Amabile, 1982) to be a creative task. In Deci's (1971) classic CET study, for example, a puzzle was used as the interesting task. However, solving a puzzle does not necessarily require a novel solution and heuristic approach and may also not meet the usefulness standard.

controlled by extrinsic motivators (p. 77). Again, this aligns with the SDT perspective that people perceive PFP to be controlling, which harms self-determination, intrinsic motivation, and thus creative performance (e.g., Deci & Ryan, 1985).

However, creativity theory began to evolve to recognize the possibility that extrinsic motivators can enhance intrinsic motivation and creative performance. Specifically, Amabile and colleagues (Amabile & Pratt, 2016; Hennessey & Amabile, 2010) proposed that, as a part of the "radically modified" creativity theory, "boosting effects" between intrinsic and extrinsic motivation may exist when intrinsic motivation is initially high and when extrinsic motivators are implemented in an autonomy-supportive work environment. Such an argument is, at the very least, logically more in line with traditional motivation theories (versus CET and SDT) that task interest and extrinsic incentives combine additively in determining performance (Latham, 2007; Vroom, 1964).[5] The results from Byron and Khazanchi's (2012) meta-analysis are consistent with the revised creativity theory argument such that creative performance-contingent reward did not undermine and indeed boosted creative performance (Hedge's g = .62 for experimental studies and .07 for field studies).

However, their meta-analysis was based on primary studies not designed to draw workplace implications (see again **Figure 1**). First, as noted earlier, these previous laboratory experiments were conducted in settings quite dissimilar from workplace settings and similarly using non-adult subject (e.g., preschool, elementary school; e.g., Boggiano & Ruble, 1979; Eisenberger & Armeli, 1997; Eisenberger & Rhoades, 2001; Lepper, Greene, & Nisbett, 1973).

[5] CET and SDT argue that intrinsic and extrinsic motivation are not additive such that total motivation is less than the sum of the two types of motivation. In contrast, the revised creativity suggests the opposite direction of the non-additive nature of intrinsic and extrinsic motivation. That is, because intrinsic and extrinsic motivation may combine *synergistically,* total motivation is greater than the sum of the two.

It is not clear that the results found in such contexts can provide meaningful implications for the workplace and/or an adult working population (Ilgen, 1986; Rynes et al., 1983; Shadish et al., 2002)[6]. Second, these primary studies defined rewards very broadly, typically using non-contingent or non-monetary rewards, such as extra course credit, teacher's verbal praise (e.g., Dodd & Ganster, 1996; Friedman, 2009; Glover & Zimmer, 1982; Jaussi & Dionne, 2003). Even in the rare case where monetary incentives were manipulated, their designs were not representative of the workplace in that they provided a piece-rate of a cent, a total payout much lower than minimum wage, or incentives without any base pay (e.g., Ariely, Gneezy, Loewenstein, & Mazar, 2009; Eisenberger & Armeli, 1997)[7]. In contrast, the current study solely focuses on performance-contingent, monetary incentives that resemble the typical incentive structures among many U.S. companies.

Third, these primary studies measured creative performance in tasks that may not be representative of creative tasks (requiring novelty and usefulness) in the workplace. Specifically, previous studies used tasks such as generating titles for short stories such as "popcorn story" where a child imagines to be a popcorn kernel in a frying pan, listing novel uses for common

[6] Researchers also suggested that results from children subjects might not be generalized to adult subjects because adults have better cognitive ability to understand and react differently to rewards (Deci et al., 1999).

[7] For example, Eisenberger and Armeli (1997) rewarded a penny or five pennies to children subjects in each of the small and large reward conditions for every correctly listed usage of a common object (e.g., paper clip, spoon, rubber band). Ariely et al. (2009) manipulated three PFP conditions (low, medium, and high) that provided $0.07, $0.70, or $7.30 (after calculating currency exchange and CPI of December 2021) for "very good" performance, and half of each amount for "good" performance, without any guaranteed base pay. In their experiment, 28 out of 87 subjects who failed to meet "good" performance did not receive any monetary reward for participating in the entire experiment (i.e., carrying out six experimental games).

physical objects, drawing pictures using circles, and creating riddles with given words (e.g., Eisenberger & Armeli, 1997; Eisenberger & Rhoades, 2001; Friedman, 2009; Kachelmeier et al., 2008). Related to this issue, in previous laboratory experiment studies, the manipulations of task autonomy, which is a key theoretical moderator addressed in the later section of this paper, were also very far from what workers experience in workplaces. For instance, it is questionable to what extent having a choice to draw either a happy face or a different object within a printed circle of 3.52 cm in diameter (Eisenberger & Armeli, 1997) can be transferred to perceived task autonomy among workers in workplaces.

A final limitation, not unique to this body of work, is the use of a between-subjects design that randomly assigns pay conditions. Although random assignment is uniquely valuable for forming equivalent groups, stopping there is not adequate to capture individual mobility, which is pervasive in the workplace. As noted, the between-subjects design precludes subjects from engaging in mobility to different pay conditions that enables a better match. Thus, it fails to incorporate the critical labor market process (i.e., the sorting effect) whereby workers sort into organizations that match better with their ability and preferences. (The sorting effect of PFP will be discussed in more detail below.) In summary, previous studies have taken important steps in understanding the (broadly defined) extrinsic rewards-creativity relationship. However, the settings in this paradigm are so different from most workplaces. Hence, the degree to which the direction and magnitude of the effect will generalize to and meaningfully inform workplace decisions on how PFP influences creative performance is open to question.

Compared to previous research on PFP and creative performance, my study uses more realistic pay designs, adult subjects, creative tasks more similar to those found in the workplace and allows mobility between pay conditions to incorporate sorting. It thus has the potential to

provide more meaningful estimate of the nature of the PFP-creative performance relationship. The incorporation of a key theoretical moderator (i.e., task autonomy in the later section of this paper) further strengthens the potential contribution.

In essence, despite the importance of creativity in organizational innovation, growth, and success and continued competing theoretical predictions regarding whether PFP will serve to help or harm creative performance in workplaces, the existing scientific evidence provides very limited guidance (Gerhart & Fang, 2015; Kim et al., 2022; Park et al., in press). To advance our understanding of how PFP influences creative performance, I first examine the effects of the absence versus presence of PFP (Hypotheses 1a and 1b) then the effects of PFP intensity as a more refined investigation (Hypotheses 2a and 2b). Based on the preceding review, I suggest two competing hypotheses. From SDT and traditional creativity theory, creative performance is higher under fixed pay than under PFP. On the other hand, as traditional motivation theories and revised creativity theory suggest, creative performance is higher under PFP than under fixed pay.

Hypothesis 1a*: Creative performance is higher under fixed pay than under PFP.*

Hypothesis 1b*: Creative performance is higher under PFP than under fixed pay.*

Incentive Effect of Different PFP Intensity on Creative Performance

A logical extension from the dichotomy of fixed pay (i.e., absence of PFP) versus PFP is to examine varying strengths of PFP, namely, PFP intensity (Gerhart, 2017; Kim et al., 2022). In the U.S. workplace, many private sector firms use PFP (Gerhart & Newman, 2020). For many, the question is not whether to use PFP, but rather what are the consequences of different levels of PFP intensity? By way of definition, high (low) PFP intensity indicates that employees receive substantially higher (lower) performance-contingent incentives from substantially higher (lower) performance (Kim et al., 2022; Milgrom & Roberts, 1992). Gerhart and Fang (2014) raised two

contrasting concerns with regard to PFP intensity. On the one hand, if PFP intensity is too low, PFP will not influence one's motivation or thus performance. This, in turn, means that, for a PFP scheme to be effective, PFP intensity should be sufficiently high. On the other hand, grounded in agency theory, they proposed that high PFP intensity may raise agency costs (i.e., adverse selection and moral hazard) as well as the risk of unintended negative consequences (Milgrom & Roberts, 1992; Prendergast, 1999).[8]

Traditional motivation theories, especially expectancy theory (Vroom, 1964) emphasize the former concern such that there should be sufficiently strong PFP intensity to prompt higher motivation and performance via the expectation that higher performance will result in higher earnings (i.e., high instrumentality). Consistent with this, Gneezy and Rustichini (2000) examined the effects of fixed pay and three levels of PFP intensity on task performance and showed that performance was highest among the mid and high PFP intensity groups, stating that "*pay enough or don't pay at all*." However, their findings were for *noninteresting* tasks (e.g., GMAT/ IQ exam) where all theories agree on a positive effect of PFP on performance. In contrast, the SDT logic would predict that higher PFP intensity can be more detrimental to performance in interesting tasks (and thus in creative tasks according to traditional creativity theory) than lower PFP intensity because a stronger PFP scheme will be perceived to be more controlling of one's motivation and behavior. Similarly, Ariely et al. (2009) argued that higher incentives drive one's "mental processes from 'automatic' to 'controlled'"(p. 452).[9]

---

[8] Another concern based on agency theory is that when strength of PFP intensity goes beyond a certain point, the advantage of stronger incentive effects may be more than offset by an agent's concern about taking on more risk (variability, uncertainty) in their stream of income.

[9] Ariely et al. (2009) found creative performance (as they measured it) was lowest when PFP intensity was highest. However, they used the "Packing Quarters" game to measure creative performance where

Kim et al. (2022) developed two measures of PFP intensity. Using the first measure (i.e., ratio of average earnings in the incentive condition to average earning in the fixed pay condition, p. 158), they showed that the relationship between PFP and performance in interesting tasks became more positive as PFP intensity increased. The effect sizes were (δ) of .62, .51, and .33 for high, medium, and low PFP intensity, respectively. However, they did not find a positive linear trend between PFP and performance in interesting tasks using the second measure (i.e., the ratio of the standard deviation of earnings in the incentive condition to the average earning in the fixed pay condition, p. 158), reporting δ = .54, .59, and .33 for high, medium, and low PFP intensity, respectively. Taken together, Kim et al. (2022) concluded there is, at a minimum, *no* negative relationship between PFP intensity and performance in interesting tasks. Consistent with Gneezy and Rustichini (2000), they found that low PFP intensity had a smaller effect size than medium, or high PFP intensity.

However, Kim et al. (2022) pointed out that their study of PFP intensity was limited in several ways due to the primary studies' research design. For instance, Kim et al. (2022) needed to operationalize PFP intensity post hoc (which had limitations) because it was not typically manipulated in the primary studies. Hence, they called for future research using primary data that explicitly measures and examines the effects of varying levels of PFP intensity. My study responds to this call by directly defining and manipulating PFP intensity in creative tasks such that I objectively measure how the same individuals would be motivated and perform differently under three levels of PFP intensity (i.e., high PFP intensity, low PFP intensity, and fixed pay). In

---

players fit nine quarter circle pieces into a squared frame. In line with the limitations of previous experiments described earlier, it is unclear to what extent such a task requires creativity and especially creativity of the sort believed to be critical to organizational performance and competitiveness in the evolving economy.

sum, although theoretical arguments from various theories as well as empirical findings from previous studies provide valuable insights, the effects of varying degrees of PFP intensity on performance, in creative tasks particularly, remains unanswered. Hence, again, I propose two competing hypotheses:

Hypothesis 2a: *Creative performance is lower as PFP intensity increases.*

Hypothesis 2b: *Creative performance is higher as PFP intensity increases.*

Moderating Effect of Task Autonomy

In hopes of helping resolve the continuing theoretical debate, I investigate the moderating effect of task autonomy on the effect of PFP on creative performance. Autonomy is defined as "the degree to which the job provides substantial freedom, independence, and discretion to the individual" (Hackman & Oldham, 1976, p. 258). As noted, revised creativity theory emphasizes the importance of an autonomy-supportive work environment to enhance creative performance (Amabile, 1993; Amabile, Conti, Coon, Lazenby, & Herron, 1996; Amabile & Pratt, 2016). Amabile and colleagues proposed the notion of *motivational synergy* that specifies the ways in which extrinsic and intrinsic motivation can be combined. Similarly, Amabile and Pratt (2016) "radically modified" early work on creativity theory to suggest that extrinsic motivators can *bolster* intrinsic motivation and thus increase creative performance when administered within an autonomous work setting where workers can decide what work to do or how to do it and feel a sense of control over one's work (Amabile, 2016).

Likewise, one interpretation of SDT also suggests the possibility that PFP can enhance intrinsic motivation (thus implying higher creative performance) under certain circumstances. Specifically, Gagné and Deci (2005) proposed the importance of autonomy "in which tangible rewards can be used so as not to be detrimental to intrinsic motivation" (p. 356), stating that

"when rewards are administered in an autonomy-supportive climate, they are less likely to undermine intrinsic motivation, and in some cases, can enhance intrinsic motivation" (p. 354). "[C]hoice, acknowledgment of feelings, and opportunities for self-direction" facilitate an autonomy-supportive climate, while "not only tangible rewards but also threats, deadlines, directives, pressured evaluations, and imposed goals" undermine it (Ryan & Deci, 2000a, p. 70). As the foundational notion of SDT is grounded in the distinction between autonomous motivation and controlled motivation where intrinsic motivation is "inherently autonomous motivation" and of the highest "quality" (Gagné & Deci, 2005; Howard et al., 2017; Ryan & Deci, 2000a, 2017), autonomy, indeed, becomes a central piece in understanding the effects of PFP on intrinsic motivation. Howard et al. (2017) stated that "the quality of one's motivation [may be] more important than quantity of motivation, meaning that the quality of one's reasons for doing something (i.e., autonomous) is more important than one's overall quantity of motivation (i.e., one's total motivation comprising both controlled and autonomous reasons)" (p. 1361). Because "autonomy involves acting with a sense of volition and having the experience of choice" (Gagné & Deci, 2005, p. 333), an autonomy-supportive climate facilitates the internalization and ultimately the integration of extrinsic motivation toward intrinsic motivation (see a continuum of motivation in Figure 1 of Gagné & Deci, 2005, p. 336 and Figure 1 of Howard et al., 2017, p. 1347). In essence, theoretical perspectives converge to anticipate a positive effect of PFP on creative performance in autonomy-supportive work settings.

Based on these arguments, two previous meta-analyses attempted to understand the effect of autonomy on the relationship between (broadly defined) extrinsic rewards and intrinsic motivation (Patall, Cooper, & Robinson, 2008) or creativity (Byron & Khazanchi, 2012). However, the two meta-analyses do not speak to the question of my study, which concerns a

moderating effect of task autonomy (whether task autonomy makes the effect of PFP on creative performance be positive). First, Patall et al. (2008) examined intrinsic motivation, not creative performance, as their dependent variable. As noted, based on the componential model of creativity (Amabile, 1983a; Amabile & Pratt, 2016), because intrinsic motivation alone cannot be the perfect predictor of creative performance, their study is limited in providing implications on a PFP-creative performance relationship. Instead, one must measure creative performance directly.

Second, Patall et al. (2008) operationalized autonomy as providing more choices, which they described as "the most obvious way to support a person's experience of autonomy" (p. 271). They found a *negative* (but close to zero) effect of having more choices on intrinsic motivation when extrinsic rewards were given (d = -.01, -.02) and a positive effect only when rewards were not provided (d = .35, .40). They suggested that this is because, when combined with extrinsic rewards, people do not perceive having more choices as pure freedom but view it as additional *controlling* tool that demands greater self-regulatory costs and, in turn, undermines intrinsic motivation. This is divergent from the interpretation of SDT by Gagné and Deci (2005) and the modified creativity theory by Amabile and Pratt (2016) that extrinsic rewards are likely perceived to be *less controlling* when autonomy is higher. On the other hand, in contrast to Patall et al. (2008), but consistent with these recent theoretical developments in SDT and revised creativity theory, Byron and Khazanchi (2012) found the positive moderating effect of giving more choices to subjects on the relationship between creativity-contingent rewards and creativity (b = .17). However, the meta-analyses by Byron and Khazanchi (2012) and Patall et al. (2008) are limited in providing insights to actual organizations about the relationship between task autonomy, PFP, and creative performance because of key limitations of the primary studies,

which I summarized earlier in **Figure 1**. These limitations, as noted, undermine the applicability of their findings to the workplace and PFP design (Kim et al., 2022).

In sum, I suggest that task autonomy could be a key to resolve the competing theoretical argument over the effects of PFP on creative performance. This is because both modified creativity theory and the at least one strand of SDT (Gagné & Deci, 2005) recognize the possibility of a positive effect of PFP on creative performance under high autonomy. Hence, I do not propose competing hypotheses here: I expect the PFP-creative performance relationship to become positive by increasing task autonomy. That is, even if there exists an overall negative relationship between PFP and creative performance (from Hypothesis 1a), the relationship is expected to be positive when task autonomy is high. Alternatively, if there already exists an overall positive relationship between PFP and creative performance (from Hypothesis 1b), it is expected to become more positive with higher task autonomy.

> ***Hypothesis 3****: Task autonomy moderates the relationship between PFP and creative performance such that the relationship becomes (more) positive when task autonomy is higher.*

Sorting Effect of PFP

Up to this point, the discussion of the effects of PFP has been focused on only the *incentive effect* of PFP, the observed performance difference between subjects assigned to and observed under PFP versus fixed pay. The incentive effect of PFP is the only mechanism that the two previous meta-analyses of the (broadly defined) extrinsic rewards-intrinsic motivation or -creative performance relationship examined (Byron & Khazanchi, 2012; Patall et al., 2008). The second mechanism by which PFP influences performance is the *sorting effect* of PFP (Cadsby et al., 2007; Gerhart & Fang, 2014; Lazear, 2000; Shaw, 2015; Trevor et al., 2012). It refers to the

performance increase observed due to a change in the composition of the workforce upon the introduction of PFP. Specifically, evidence to date suggests that less productive employees are more likely to leave the firm after introduction of a PFP scheme because they anticipate their low performance will result in their pay being lower, in absolute or relative terms. In turn, they will be replaced by more productive employees who will fare better under PFP.

Lazear (2000) found in his influential paper that the introduction of a piece-rate (individual incentive) scheme in an automobile windshield repair company enhanced productivity by 44%. Of the total increase, one-half was explained by the incentive effect (i.e., a 22% productivity increase among the same pool of employees who were present in the company both before and after the change), while the remaining half was attributed to the sorting effect (i.e., a 22% productivity increase due to less productive employees leaving and being replaced by more productive employees). His argument is largely consistent with the Attraction-Selection-Attrition (ASA) framework (Schneider, 1987) in that pay schemes influence matching processes between organizations and employees via attraction and selection (Bretz, Ash, & Dreher, 1989; Cable & Judge, 1994) and retention (Shaw, Dineen, Fang, & Vellella, 2009; Trevor, Gerhart, & Boudreau, 1997; Trevor et al., 2012).

Importantly, however, Lazear (2000) did not study creative performance, the focus of my study. He also did not examine the role of task autonomy, which as I have described, theory specifies as a key moderator of the PFP-creative performance relationship. Further, whereas Lazear (2000) studied a single PFP intensity (versus fixed pay). My study uses two levels of PFP intensity, providing us with the opportunity to better calibrate how varying PFP intensity influences sorting.

It is crucial, as noted earlier, to realize that the sorting effect can only be observed where subjects/employees are observed over time and have the opportunity to move to a different level of PFP intensity (including none—fixed pay). Considering the large size of the sorting effect found by Lazear (2000) and the very large amount of economy-wide worker mobility described earlier (JOLTS, 2020), any study design (as is the case in a purely between-subjects design) that exclude the sorting effect may miss a significant part of the impact of PFP. Indeed, because the sorting effect indicates that organizations (and/or PFP schemes) and employees are not randomly assigned, any study that does not allow/incorporate worker mobility/pay choice (the sorting effect) is at greater risk of finding a detrimental effect of PFP that is an artifact of such a design choice. That is, even if the undermining effect of PFP on intrinsic motivation (and by implication, creative performance) exists as SDT argues in randomized experiments, which by definition produces a substantial number of mismatches, a design that allows the sorting effect of PFP to operate, resembling the actual labor market, may weaken, or even eliminate, adverse effects of PFP to the degree those who prefer PFP are able to sort into PFP and those who prefer fixed pay or low PFP intensity are able to sort into that organization (Fang & Gerhart, 2012). However, to my knowledge, none of the previous studies on creativity, especially grounded in SDT, have incorporated this important sorting mechanism, despite its potential to change observed effects. Instead, they focused solely on the incentive effect by randomly assigning different pay conditions to subjects and not allowing subjects to choose their preferred pay conditions, as opposed to what workers do in the labor market.

Cadsby et al. (2007) is the only laboratory experiment that examined the sorting of PFP in addition to the incentive effect of PFP on performance. However, as noted below, their study, while serving as a valuable model, does not speak to creative performance. It also uses a form of

piece-rate that is, in fact, rare in the workplace. They found that individuals make decisions regarding whether to stay or leave under PFP versus fixed pay depending on their abilities and preferences. Specifically, using an anagram task, they showed that more capable and risk tolerant subjects were more likely to sort into PFP than fixed pay, whereas less capable and risk averse subjects were more likely to sort into fixed pay than PFP. However, it is important to note that like Lazear (2000), Cadsby et al. (2007) used a *non-creative,* and arguably *noninteresting* task where all theories agree in predicting a positive effect of PFP on performance. In contrast, my focus is on how PFP influences creative performance, where, as we have seen, there is no theoretical consensus on the effect of PFP. The two studies also used a piece-rate scheme (a form of individual incentive), where performance is measured objectively and the payoff scheme (money earned per piece) is based on an objective formula. In contrast, the connection, objective and perceived, is likely to be weaker in contemporary organizations for two reasons. First, they more commonly use a subjective evaluation (rating) of performance by supervisors, sometimes also colleagues or subordinates (Gerhart & Fang, 2015; Gerhart, Rynes, & Fulmer, 2009). Second, in organizations, the payoff to a performance rating is also typically subjective (entirely or in part) in the workplace (Gerhart & Newman, 2020). Thus, we still do not know how the sorting mechanism will operate in tasks requiring creativity and in tasks where PFP is based on performance being evaluated and rewarded subjectively, which is more typical of the workplace and may reduce perceived PFP.

To understand the sorting mechanism in creative tasks, I consider how self-assessment of creative performance, in addition to actual creative performance evaluated by others, influences one's sorting behaviors. An implicit assumption is that individuals are effective in sorting in and out of different pay schemes according to their expected performance and thus expected

earnings. However, it is important to recognize that one's perception of their performance, which drives sorting decisions, may not necessarily be congruent with one's actual performance. That is, an employee may not have an accurate knowledge about her ability and/or the performance evaluation system such that she may not be able to accurately predict her evaluated performance (and thus her resulting pay). To the degree this incongruence/lack of accuracy exists, the sorting mechanism observed in previous research cannot take place.

I expect such incongruence between (evaluated) actual performance and (self-assessed) expected future performance to become more substantial given three characteristics of the current study. First, unlike Cadsby et al. (2007) and Lazear (2000), my dependent variable is creative performance. By its nature, creativity requires greater cognitive ability, divergent thinking, a more heuristic approach, and domain-specific skills. This may increase the uncertainty on *how* to best carry out tasks to achieve high (evaluated) actual creative performance compared to the case where tasks are non-creative, simple, and well-defined (Amabile, 1982, 1983a; Baer, 2010; Pretz & McCollum, 2014). Second, performance in non-creative, simple, and well-defined tasks can be objectively quantified, as well as easily observable and evaluated. Creative tasks, in contrast, involve greater uncertainty because of the subjectivity of performance evaluation and the less clear link (Vroom, 1964) between (a) one's effort and such subjectively evaluated creative performance (expectancy) and (b) one's evaluated creative performance and pay (instrumentality). For instance, predicting the degree to which your design of a new car will be evaluated (and rewarded) as creative is more uncertain than counting the number of car windshields you repaired (Lazear, 2000) or the number of anagrams you solved (Cadsby et al., 2007) and applying an objective payoff formula. Third, I use a relative performance rating system, which likely adds another layer of uncertainty in predicting one's

own creative performance, compared to using an absolute performance scale (Gibbons & Murphy, 1999; Holmstrom, 1982; Lazear & Rosen, 1981). This is because workers need to gauge not only how well they performed but also where their performance levels would stand among peers/competitors. I chose these three design characteristics because they are typical of workplace settings.

Even though creative tasks may cause greater uncertainty in predicting one's own creative performance, research has consistently shown a positive relationship between self-perceived creativity and divergent thinking evaluated by others correlations in the range of .15 to .27 (e.g., Batey, Furnham, & Safiullina, 2010; Furnham, Zhang, & Chamorro-Premuzic, 2005; Grohman, Wodniecka, & Klusak, 2006; Pretz & McCollum, 2014), (Divergent thinking, in contrast to convergent thinking, has long been regarded by most to be a key predictor of creativity (Guilford, 1967; Runco & Acar, 2012))[10]. That self-assessments and assessments by others of performance are not interchangeable is well-known (Garden, 1991; Park, Chun, & Lee, 2016; Podsakoff & Philip, 1986) and an important aspect of the workplace.[11]

In sum, my goal is to examine the degree to which self-assessed and actual (as measured by independent raters) creative performance influences the choice of PFP versus fixed pay. Specifically, because people make sorting decisions based on their self-assessed creative performance and thus their projected earnings, I predict that people are more likely to choose a stronger PFP scheme if they self-assess their creative performance to be higher. As part of this, I examine the degree to which actual creative performance influences choice of PFP versus fixed

[10] Note that some scholars criticize using divergent thinking to measure creativity, stating that it measures "only the most trivial aspects of creativity" (Sternberg, 1985).

[11] In work settings, self-assessed creative performance is expected to become less uncertain and much closer to actual creative performance as workers gain enough work experience.

pay via self-assessed creative performance. That is, given the likelihood that workers are not able to perfectly predict their actual creative performance, but that there is some meaningful relationship between it and their self-assessed creative performance, it is of interest to examine not only the degree to which self-assessed creative performance is associated with sorting into PFP, but also its interplay with actual creative performance, the focus in previous studies of sorting (which used non-creative tasks as well as objective and non-relative performance evaluations). Thus, I hypothesize:

Hypothesis 4*: The higher (lower) actual creative performance, the more likely the choice of higher (lower) PFP intensity.*

Hypothesis 5*: The higher (lower) self-assessed creative performance, the more likely the choice of higher (lower) PFP intensity.*

Hypothesis 6*: The influence of actual creative performance on choice of PFP intensity is mediated by self-assessed creative performance.*

Finally, I predict that the gap between creative performance under PFP and that under fixed pay is greater when people are allowed to sort into their preferred pay conditions than when they have no choice but to work under randomly assigned pay conditions. That is, under self-selection, both the incentive and sorting effects of PFP operate to enhance creative performance among those who prefer and sort into PFP, whereas under random assignment, only the incentive effect operates. The current study's laboratory experiment design that incorporates both random assignment and self-selection of pay conditions provides the necessary design to investigate to what extent the sorting effect of PFP exists by decomposing it from the incentive and total effect of PFP.

Hypothesis 7*: The effect of PFP on creative performance is greater when PFP is self-selected (i.e., incentive and sorting effects are combined) than when it is randomly assigned (i.e., only incentive effect is present).*

Methods

Sample, Design, and Procedures

Participants were recruited from the *Behavioral Research Insights Through Experiments (BRITE) Lab* at the University of Wisconsin-Madison that manages a large pool of participants and administers behavioral and experimental social science research (IRB Approval No. 2021-0628). Potential participants received an email that invites participation in the experiment with brief information on the study (see **Appendix A**). The experiment was held virtually so that participants used their personal computers or the computer labs on campus. As informed in the recruiting email, the entire experiment took approximately 1 hour, including the general study instruction and study consent, four task sessions for up to 10 minutes each, and a series of debriefing questions at the end. The experiment consists of two between-subjects (two creative tasks X two task autonomy levels) and one within-subjects (three pay random assignment sessions + one pay choice session) factor (see **Figure 4**).

Creative Tasks

Participants were randomly assigned to one of the two creative tasks: (1) *slogan task*: creating advertising slogans for innovative products for client firms as copywriters at an advertising company (see **Appendix B** for the detailed task instruction) and (2) *article task*: writing short magazine articles on familiar topics after conducting web search as magazine journalists (see **Appendix C** for the detailed task instruction). Unlike simple, non-work-like tasks often used in previous studies, the two creative tasks in this study were designed to

simulate real work-like situations. Furthermore, although the two creative tasks were found to show a similar degree of creativity, enjoyment, interestingness, and satisfaction (see the Results section), they may require different dimensions of work skills, such as imagination and divergent thinking for the slogan task versus web search and writing skills for the article task. Hence, by including the two creative tasks, the current study can provide implications to a broader range of creative works. Lastly, the task design allows generating multiple versions by using different sets of products for slogan and topics for article with little learning effect over task sessions (see **Appendix D** for the task topics), providing a setting to examine the incentive and sorting effects of different pay conditions in a within-subjects manner.

Task Autonomy

The other between-subjects manipulation is *task autonomy*. Participants were randomly assigned to either high or low task autonomy. Under the *high* task autonomy condition, participants were given three topics, and they had the freedom to choose which topic to work on, whereas under the *low* task autonomy condition, such options were not offered. This manipulation aims to influence perceived task autonomy regarding "what work to do" (Amabile, Conti, Coon, Lazenby, & Herron, 1996, p. 1166), as Patall et al. (2008) argued that the provision of choice is "the most obvious way to support a person's experience of autonomy" (p. 271). The two task autonomy conditions were randomly assigned across all participants. For instance, as **Figure 4** depicts, Participant #1 worked on the slogan task under high task autonomy throughout the four task sessions, while Participant #2 performed the same task under low task autonomy. Participant #3 conducted the article task under high task autonomy throughout the entire experiment, and Participant #4 worked on the same task under low task autonomy.

Pay for Performance (PFP)

The four task sessions consist of *three randomly assigned pay conditions* and *one pay choice condition* at the end as a within-subjects manipulation (see **Figure 4**). That is, in each of the first three task sessions, participants worked under either fixed pay, low PFP intensity, or high PFP intensity conditions in random order to address any potential order effects over multiple task sessions. Then, to examine the sorting mechanism, the fourth task session allowed participants to choose the pay condition they preferred to work under among the three pay conditions they already experienced in the first three preceding task sessions. The three pay conditions were chosen to correspond to those found in workplace settings. First, the overall pay level was set between $12.10 and $24.50 (between $13.09 and $28.70, once the earnings from the lottery were included; see the Supplementary Analyses section of the measure for *risk aversion*) such that it was commensurate with or substantially higher than what participants could typically earn from on-campus jobs[12].

Second, I benchmarked the results from large-scale surveys of U.S. companies in designing performance rating cutoffs (Gerhart & Newman, 2020) and incentive rates (WorldatWork, 2018). Hence, the three pay conditions aim to manipulate sufficient pay differentials across low, average, and high performance ratings under the two PFP conditions as well as strong enough incentives under the two PFP conditions compared to the fixed pay condition. **Table 1** and **Figure 5** show a summary of three pay conditions, and **Appendix E** compares the three pay conditions in more detail. Under the fixed pay condition, participants received $3.90 regardless of how their creative performance was evaluated. The low PFP intensity condition had a base pay of $3.20 and an incentive opportunity of $2.00 for the top 30% of creative performance; $1.00 for the next 50%; and no incentive for the bottom 20%. In

[12] Hourly pay for a campus job ranges from $10 to $15 posted on https://studentjobs.wisc.edu.

comparison, under the high PFP intensity condition, base pay was lower ($2.50), but an incentive opportunity was higher ($5.20 for top 30% of creative performance; $2.10 for next 50%; and no incentive for the bottom 20%)[13]. The ratio of the potential earnings (base pay + incentive) by a high (Top 30%) performer to an average (Next 50%) performer is 1.24 (= $5.20/$4.20) for the low PFP intensity condition and 1.67 (= $7.70/$4.60) for the high PFP intensity condition. Such ratios correspond to ratios reported by publicly traded U.S. firms in which the lower end of PFP intensity ranges between 1.10 to 1.24 versus a range at the upper end of 1.50 to 1.74 (WorldatWork, 2018). Unlike the fixed pay condition, the two PFP conditions introduce pay at risk because their base pay is lower than the fixed pay, meaning that participants can earn less under the two PFP conditions if their performance is low. However, the two PFP conditions also provide a risk premium such that participants can earn more under the two PFP conditions than in the fixed pay condition by performing well. In the high PFP intensity condition, both incentive intensity and risk are higher than in the low PFP intensity condition.

Measures

Creative performance for the two creative tasks was evaluated using a consensual assessment technique developed by Amabile (1982), which is commonly used in creativity research (e.g., Binnewies, Ohly, & Sonnentag, 2007; Shalley, 1995; Zhou, 1998; Zhou & Oldham, 2001). Two fourth-year Ph.D. students in Management served as independent judges who were blind to the research question, hypotheses, and experimental conditions. They were informed the definition of creativity as well as the evaluation standards and instructions given to participants, which emphasized the importance of novelty and usefulness in generating creative

[13] Payouts and performance rating cutoffs were revised after the pilot test to increase the overall expected earning and PFP intensity.

ideas. They rated each slogan (or article) based on its novelty, usefulness, and overall creativity (novelty and usefulness combined) on a six-point scale (1 = Not creative at all, 6 = Extremely creative). The average of the overall creativity between the two judges was used, and interrater reliability was .75 (ICC(3,1)) and .86 (ICC(3,2)).

To measure participants' *self-assessment of creative performance*, I asked, "how would you assess your performance? If there are a total of 100 participants (including you) in this study, what will your rank be in each session? (Rank #1: Best performer … Rank #100: Bottom performer)" The variable used in the analyses was reverse coded for easier interpretation such that the higher the number, the better the creative performance evaluated by each participant.

In the debriefing session, I asked a series of questions about how participants perceived the study manipulations. First, participants were asked to indicate the degree to which they felt *creativity was needed* in the task they worked on (1 = Not at all, 5 = A great deal). In addition, to understand how participants experienced the creative tasks, I asked a series of questions asking the degree to which they felt the tasks were interesting versus boring (1 = Very bored, 5 = Very interested), enjoyable and fun (1 = Not at all, 5 = A great deal), satisfying (1 = Very unsatisfied, 5 = Very satisfied), and challenging (1 = Not at all, 5 = A great deal). The *task autonomy* manipulation was checked by asking participants the degree to which they felt they had control over what work to do (1 = Not at all (no control), 5 = A great deal (high control)). In addition, the effectiveness of the *pay condition* was assessed by asking participants to indicate the degree to which they felt their pay was dependent on how they performed for each session (1 = Not dependent at all, 3 = Very dependent).

Pilot Study

The purposes of the pilot study were twofold. First, it was to test whether participants understood the instructions/manipulations. Second, because two creative tasks were newly developed for this study, the pilot study sought to investigate the degree to which participants perceived them to be similar or different from the two noninteresting tasks (described in the following section) in terms of creativity and task interest.

Sample, Design, and Procedures

A total of 16 undergraduate and graduate students at the University of Wisconsin-Madison were recruited from the BRITE Lab, in which the main experiment was administered. Participants' age ranged from 19 to 28 and, on average, 22.6. 13 out of 16 participants were female. Participants were from various majors, including biology, social work, mass communication, business, economics, medicine, history, and anthropology.

The entire experiment was conducted virtually for approximately 1 hour, including the general study instruction and study consent, four task sessions that each lasted up to 10 minutes, and a series of debriefing questions at the end. The four task sessions in the pilot test were designed a little differently from the main experiment. One difference is the pilot study included two noninteresting tasks. They were necessary as a baseline comparison to establish whether the two creative tasks designed for this study indeed involve more creativity and are interesting. Participants performed one of the two noninteresting tasks (*inventory task* vs. *reference task;* see below for descriptions) twice each under fixed pay and PFP; and one of the two creative tasks (*slogan task* vs. *article task*) twice each under fixed pay and PFP. Then, participants were randomly assigned to either high or low task autonomy, as suggested in the main experiment. For instance, Participant #1 had high task autonomy and conducted two sessions of the inventory task, one under PFP and one under fixed pay, and two sessions of the slogan task, one under PFP

and one under fixed pay. Meanwhile, Participant #2 had low task autonomy and carried out two sessions of the reference task each under PFP and fixed pay, and two sessions of the article task each under PFP and fixed pay. See **Appendix F** for the summary of random assignments.

The two noninteresting tasks were designed in a way that creativity, task interest, or enjoyment was not involved (Cerasoli et al., 2014; Gagné & Deci, 2005). In the inventory task, participants sorted the randomly listed inventories of a food processing company in ascending order of alphabet and dates (see **Appendix G** for the details). Under the PFP condition, base pay was $2.70, and an incentive was $0.10 for every correctly sorted inventory item, which sums up to a maximum of $7.70 if participants successfully sort all the 50 items provided. Participants received $3.80 regardless of the number of correctly sorted items. On the other hand, in the reference task, participants were asked to organize reference lists of academic articles following the given citation rules (see **Appendix H** for the details). Under the PFP condition, they earned a base pay of $2.70 and an incentive of $0.40 for every correctly organized reference article such that they were able to earn up to $7.50 if they correctly organized all 12 articles; and under the fixed pay condition, they received $3.80 regardless of the number of successfully organized reference articles. Finally, in the two creative tasks, the PFP condition had a base pay of $2.70 and an incentive opportunity of $4.30 for the top 20% of creative performance; $1.60 for the next 50%; and no incentive for the bottom 30%. Under the fixed pay condition, participants received $3.80 regardless of their creative performance.

Measures

Creative performance for the two creative tasks was evaluated in the same way suggested in the main experiment. The interrater reliability was .70 (ICC(3,1)) and .81 (ICC(3,2)).

Ordinary performance for noninteresting tasks was measured by counting the accurately completed items. The other measures were the same as the ones in the main experiment.

Descriptive Statistics and Manipulation Checks (Pilot Study)

Table 2 shows the means, standard deviations, number of observations, and intercorrelations among the variables. **Table 3** provides the information of most interest, which is the effectiveness of the three main experimental manipulations—task types, task autonomy, and pay conditions. First, participants perceived that the two creative tasks were significantly higher in need for creativity, interestingness, enjoyment and fun, and satisfaction compared to the two noninteresting tasks. The right most column in **Table 3A** shows that variance explained by task (creative versus noninteresting) ranged from 23% to 88%. However, perceived challenge did not differ between creative and noninteresting tasks. The second key manipulation, task autonomy, was also effective. **Table 3B** shows participants in the high task autonomy condition perceived significantly higher personal control over what work to perform ($t(14) = 3.07$, $p < .05$, $r_{pb}^2 = .40$). Finally, as shown in **Table 3C**, participants perceived their pay as more dependent on how they performed under the PFP condition than under the fixed pay condition ($t(62) = 5.40$, $p < .05$, $r_{pb}^2 = .32$). In sum, the results from the pilot test confirmed that the two creative tasks developed for this study were indeed appropriate to be used to examine creative performance and that the experimental manipulations of task autonomy and pay conditions were also effective. The power analysis of the pilot study indicated that a sample size of 40 to 82 for each of the slogan and the article tasks would be needed to detect statistical significance of effects. Hence, I aimed to recruit 100 participants for each of the two tasks (i.e., 200 participants in total).

Descriptive Statistics and Manipulation Checks (Main Experiment)

A total of 218 undergraduate and graduate students participated in the main experiment. Participants' age ranged from 18 to 63, mean of 22.9. 161 participants were female, and 4 participants preferred not to identify their gender. Participants were from diverse educational backgrounds, including nursing, biochemistry, linguistics, business, economics, medicine, communication, engineering, kinesiology, psychology, and so forth.

Table 4 shows the means, standard deviations, number of observations, and intercorrelations among the variables. **Table 5** provides data on the effectiveness of the three major experimental manipulations. The two creative tasks showed a relatively high need for creativity (means of 4.28 and 4.04 on a 5-point scale). Further, means in **Table 5A** for interestingness, enjoyment and fun, and satisfaction are all higher than were the corresponding means for the noninteresting tasks in the pilot study in **Table 3A**. Finally, as in the pilot study, challengingness is not perceived as high. Supporting the effectiveness of the task autonomy manipulation, as **Table 5B** shows, participants in the high task autonomy condition perceived significantly higher personal control over what work to perform ($t(216) = 14.75$, $p < .01$, $r_{pb}^2 = .50$). Finally, as reported in **Table 5C**, successful manipulation of the three pay conditions was also confirmed with participants perceiving their pay as more dependent on how they perform under high PFP intensity condition than low PFP intensity condition ($t(217) = 10.54$, $p < .01$, $r_{pb}^2 = .17$) and fixed pay condition ($t(217) = 28.33$, $p < .01$, $r_{pb}^2 = .66$).

Hypotheses Testing

I used a repeated-measures ANOVA to test the competing hypotheses that creative performance is higher under fixed pay (Hypothesis 1a) versus higher under PFP (Hypothesis 1b). As Model 4 of **Table 6** shows, creative performance was higher under the two PFP conditions (high and low PFP intensity combined) than fixed pay, supporting Hypothesis 1b, $F(1, 435) =$

107.82, $p < .001$, $\omega^2 = .20$. (See **Appendix I** for the means and standard deviations of creative performance under different experimental conditions and **Appendix J** for detailed repeated-measures ANOVA results.) Creative performance was significantly higher by 29% under PFP ($M = 3.38$) than fixed pay ($M = 2.63$). This was when participants were randomly assigned to each pay condition, examining the incentive effects of PFP.[14]

In terms of testing whether creative performance is lower (Hypothesis 2a) versus higher (Hypothesis 2b) as PFP intensity increases, I compared creative performance under high PFP intensity, low PFP intensity, and fixed pay. Models 1, 2, and 3 of **Table 6** indicate these differences are statistically significant. Creative performance increased by 23%, moving from fixed pay to low PFP intensity ($M = 3.24$) and by 34% moving from fixed pay to high PFP intensity ($M = 3.53$). In sum, Hypothesis 2b is supported such that the stronger PFP intensity, the greater the positive effect of PFP on creative performance.

Next, I tested Hypothesis 3 that task autonomy moderates the relationship between PFP and creative performance. I found a statistically significant interaction between task autonomy and the three pay conditions on creative performance, both using the two PFP conditions combined and the two PFP conditions separately. Specifically, **Table 7A** and **Figure 6A** show that, under high task autonomy, creative performance increased by 38% moving from fixed pay

[14] To examine whether effects of PFP differ across the two creative tasks, I conducted a split-plot ANOVA of PFP and the two creative tasks (i.e., the *article* and *slogan* tasks) on creative performance. The PFP-by-task type interaction was statistically significant ($F(2, 432) = 3.06$, $p < .05$) as were the two main-effects for PFP ($F(2, 432) = 64.05$, $p < .01$) and task type ($F(2, 432) = 4.94$, $p < .05$). The estimates of variance explained for the interaction ($\omega^2 = .01$) and task type ($\omega^2 = .02$) were much smaller than that of the PFP ($\omega^2 = .22$). The pairwise contrast further showed that creative performance under the article task ($M = 3.78$) was significantly higher by 14% than the slogan task ($M = 3.32$) only at high PFP intensity ($t(216) = 3.08$, $p < .01$).

(M = 2.61) to PFP (M = 3.14), whereas, under low task autonomy, it increased less, by 20%, moving from fixed pay (M = 2.64) to PFP (M = 3.63), $F(1, 434) = 10.01$, $p < .01$, $\omega^2 = .02$. Further, I found that task autonomy did not significantly increase creative performance under fixed pay, a difference of 1%. **Figure 6B** and **Table 7B** report the same analyses, but with high and low PFP intensity conditions being examined separately. Again, I found a statistically significant interaction term between task autonomy and pay conditions, $F(2, 432) = 8.84$, $p < .001$, $\omega^2 = .03$. Specifically, mean creative performance differences between fixed pay and high PFP intensity are significantly different across high and low task autonomy (i.e., $M_{high\ PFP,high\ autonomy} - M_{fixed\ pay,high\ autonomy} \neq M_{high\ PFP,low\ autonomy} - M_{fixed\ pay,low\ autonomy}$), $F(1, 432) = 17.16$, $p < .001$). While creative performance increased by 47% under high task autonomy moving from fixed pay (M = 2.64) to high PFP intensity (M = 3.88), it increased less, by 22%, under low task autonomy moving from fixed pay (M = 2.61) to high PFP intensity (M = 3.18). It is noteworthy that increasing PFP intensity from low (M = 3.11) to high (M = 3.18) enhanced creative performance only by 2%, when task autonomy was low, whereas increasing PFP intensity from low (M = 3.37) to high (M = 3.88) enhanced creative performance by 15% when task autonomy was high, indicating a significant difference in the effects of moving from low to high PFP intensity as a function of task autonomy (i.e., $M_{high\ PFP,high\ autonomy} - M_{low\ PFP,high\ autonomy} \neq M_{high\ PFP,low\ autonomy} - M_{low\ PFP,low\ autonomy}$; $F(1, 432) = 7.26$, $p < .01$).

To understand the sorting effect, I examined whether participants were more likely to choose high PFP intensity over fixed pay when their actual (Hypothesis 4) and self-assessed (Hypothesis 5) creative performance was higher and whether the effect of actual creative performance on pay choice would be mediated by self-assessed creative performance

(Hypothesis 6). To begin, I found that in the pay self-selection session (i.e., session 4), out of 218 participants, 48 participants chose fixed pay, 58 participants low PFP intensity, and 112 participants high PFP intensity. Thus, 78% of the participants preferred PFP to fixed pay. I also found a positive and statistically significant correlation between actual creative performance and self-assessed creative performance ($r = .19$, $p < .01$).[15]

To test Hypothesis 4 where I predicted that the higher actual creative performance, the more likely sorting into higher PFP intensity, I conducted logistic (**Table 8A**) and multinomial logistic (**Table 8B**) regression analyses to estimate the probability of sorting into PFP versus fixed pay predicted by actual creative performance. As shown in Model 1 of **Table 8A**, as expected, the coefficient on actual creative performance is positive, indicating higher performers are more likely to sort into PFP. **Table 8B** shows the same analyses, but with high and low PFP intensity conditions examined separately, reporting relative risk ratio[16] in addition to coefficients of selecting high and low PFP intensity over fixed pay. As **Figure 7** shows, I found that going from one standard deviation below the mean to one standard deviation above the mean in actual creative performance increased the probability of sorting into high PFP intensity by 46 percentage points (from 28% to 74%)[17], decreased the probability of sorting into low PFP

[15] This correlation is likely higher using ordinary, objective performance, where outputs are easily observed or assessed (and rewarded based on a formula) and performance is not relative (e.g., counting the number of repaired automobile windshields as in Lazear (2000).

[16] The probability of self-selecting PFP over the probability of self-selecting fixed pay (i.e., base category) is expressed in the relative risk ratio. In contrast to reporting ratios of odds for logistic regression analysis where only two categories of dependent variables are considered, reporting relative risk ratio by exponentiating coefficients is accurate when more than two categories of dependent variables are examined simultaneously (STATA, 2016).

[17] Results remain consistent when the probabilities of sorting into high PFP intensity were examined at 25th, 50th, and 75th percentiles of actual creative performance, which were 33%, 54% and 64%.

intensity by 8 percentage points (from 30% to 22%), and decreased the probability of sorting into fixed pay by 38 percentage points (from 42% to 4%), supporting Hypothesis 4.

Next, as I anticipated in Hypothesis 5, I found that self-assessed creative performance (also) significantly predicts the probability of sorting into PFP when the high and low PFP intensity conditions are combined (Model 2 of **Table 8A**) as well as when they are separately examined (Model 2 of **Table 8B**). **Figure 8** shows the probability of sorting into each pay condition predicted by self-assessed creative performance. Supporting Hypothesis 5, going from one standard deviation below the mean to one standard deviation above the mean in self-assessed creative performance increased the probability of choosing high PFP intensity by 78 percentage points (from 9% to 87%), decreased the probability of choosing low PFP intensity by 28 percentage points (from 40% to 12%), and decreased the probability of choosing fixed pay by 50 percentage points (from 51% to 1%). When the high and low PFP intensity conditions were combined, moving from one standard deviation below the mean to one standard deviation above the mean in self-assessed creative performance increased the probability of sorting into PFP by 47 percentage points (from 52% to 99%) but decreased the probability of sorting into fixed pay by 47 percentage points (from 48% to 1%). Note also that Figure 7 and Figure 8 have very similar patterns, indicating very similar results whether using actual creative performance or self-assessed creative performance, but with stronger results using the latter.

In Hypothesis 6, I proposed that self-assessed creative performance would mediate the relationship between actual creative performance and pay choice. As indicated by the model chi-square ($\mathcal{X}^2$) in **Table 8**, the predictive power of actual creative performance, although substantial, is as expected not as high as that of self-assessed creative performance. **Table 8A** and **Table 8B** show the coefficients of actual creative performance from Model 1 predicting the

probability of sorting into PFP in comparison with sorting into fixed pay decreased in Model 3 by including self-assessed creative performance. Thus, the results indicate that self-assessed creative performance partially mediates the relationship between actual creative performance and sorting into PFP, supporting Hypothesis 6. Although participants could only draw an imperfect inference about their actual creative performance, it did influence their sorting decisions and did so to some degree via their self-assessed creative performance.

Finally, I predicted that the effect of PFP on creative performance would be greater when participants were allowed to sort into their preferred pay conditions than when they had to work under randomly assigned pay conditions (Hypothesis 7). I tested this first by estimating and comparing the effects of PFP on creative performance each under random assignment (i.e., sessions 1 through 3) and self-selection (i.e., session 4) conditions. **Table 9** shows the number of observations, mean creative performance by pay condition and task sessions, and the percentage differences in creative performance between PFP and fixed pay using both the number of observation-weighted and -unweighted forms. The weighted calculation addressed the unequal number of observations across pay conditions in the self-selection session such that more participants selected the two PFP conditions (especially high PFP intensity) than fixed pay. As predicted, the effect of PFP was greater when sorting was possible. Specifically, once the number of observations was accounted for, creative performance was higher by 292% among those who sorted into high PFP intensity than fixed pay, whereas I observed a 34% increase in creative performance when high PFP intensity and fixed pay were randomly assigned.

Next, to decompose and estimate the magnitude of the sorting effect of PFP from the total effect of PFP, as shown in the **Table 9**, I subtracted the effect of PFP under random assignment sessions from that under the pay self-selection session, which yielded the sorting effect of PFP

alone. Then, I divided it by the effect of PFP under the pay self-selection session, which represents the total effect of PFP where the incentive and sorting effects are combined. I found that about one-half of the total effect of high PFP intensity in increasing creative performance was attributed to the sorting effect with using the unweighted estimates. The ratios of the sorting effect to the total effect of PFP were greater once the unequal number of observations was accounted for. Specifically, 88% of the creative performance increase under high PFP intensity versus fixed pay was due to the sorting effect, 65% under low PFP intensity, and 94% under PFP where high and low PFP intensity conditions were combined. In essence, the results show that allowing the self-selection of pay conditions, analogous to the actual labor market dynamics where workers change jobs based on their ability and preferences, has a considerable impact in increasing creative performance. Hence, my study demonstrates the importance of incorporating the sorting mechanism in studying not only performance generally, but creative performance specifically. Using a study design that does not incorporate sorting will likely substantially underestimate the positive effect of PFP in workplace settings where sorting is pervasive.

Supplementary Analyses

Incentive Effects among Intrinsically Motivated Participants

As another way to examine Hypotheses 1a versus 1b, I compared creative performance under fixed pay versus PFP (or higher PFP intensity) among those who reported being intrinsically motivated to participate in this study. If the SDT argument that the detrimental effect of PFP is greater when intrinsic motivation is initially higher holds, creative performance should be lower under PFP than fixed pay among intrinsically motivated participants (e.g., Deci, 1972; Ryan & Deci, 2000b). To investigate this, I asked participants to respond to an open-ended question, “why did you decide to participate in this study?” after reading the overall experiment

instruction about the task, expected pay range, and study duration and signed the consent form. I coded the written responses into four categories: (1) extrinsically motivated; (2) intrinsically motivated; (3) both extrinsically and intrinsically motivated; and (4) uncategorizable responses. See **Table 10** for example responses and frequency for each category.

I found that, supporting Hypothesis 1b but not supporting the theoretical argument by SDT, intrinsically motivated participants showed higher creative performance under PFP (M = 3.27) than fixed pay (M = 2.41), a difference of 36%, $F(1, 119) = 33.53$, $p < .001$, $\omega^2 = .21$. This implies that intrinsic motivation and extrinsic motivation combine additively to enhance creative performance because it was higher when intrinsically motivated participants were paid by PFP (i.e., when both intrinsic motivation and extrinsic motivation existed) than under fixed pay (i.e., when only intrinsic motivation existed). Further, the creative performance increase of 36% under PFP among those reporting initial intrinsic motivation is slightly higher than the creative performance increase of 29% under PFP for the entire sample.

Examination of Mis-sorted Participants

I investigated whether self-assessed creative performance and one's risk aversion were associated with some of participants' mis-sorting decisions. There were 114 out of 218 participants who mis-sorted themselves such that self-selected pay conditions did not match with their actual creative performance. For instance, top performers who ranked in the top 30% of creative performance would have earned $7.70 if they correctly sorted into high PFP intensity condition. That compares to 51 top performers who did so (see top left gray-shaded cell in **Table 11)**. However, as shown in **Table 11**, 16 of the top performers earned only $5.20 (i.e., loss of $2.50) by choosing low PFP intensity, and four of them earned only $3.90 (i.e., loss of $3.80) by choosing fixed pay. Similarly, participants who ranked in the next 50% of creative performance

would have earned $4.20 if they correctly sorted into low PFP intensity. However, 16 of them mis-sorted into fixed pay, earning only $3.90 (i.e., loss of $.30), whereas 49 of them mis-sorted into high PFP intensity, earning $4.60 (i.e., gain of $.40). Finally, low performers who ranked in the bottom 20% could have earned $3.90 if they corrected sorted into fixed pay. However, 17 of them who mis-sorted into low PFP intensity lost $.70, and 12 of them who mis-sorted into high PFP intensity lost $1.40.

First, I examined whether mis-sorting decisions were driven by one's self-assessment of creative performance. As shown in the earlier section, because self-assessed creative performance mediates the relationship between actual creative performance and sorting into PFP versus fixed pay, participants might mis-sorted themselves if they under-assessed or over-assessed their creative performance. Specifically, participants might have sorted into weaker PFP intensity conditions because they under-assessed their creative performance compared to their relatively high actual creative performance. They are shown in the three cells above the gray-shaded diagonal in **Table 11**. I found that these 36 participants significantly under-assessed their creative performance (M = 40.64, SD = 26.17) compared to the rest of 182 participants by 51% (M = 61.24, SD = 27.52) and 78 participants in the three cells below the diagonal who sorted into stronger PFP intensity conditions than their relatively low (actual) creative performance by 67% (M = 67.86, SD = 20.66).

Second, regarding risk aversion, PFP introduces risk in projected earnings (Brown & Huber, 1992), where risk is defined as "variance (lack of stability) of income and/or the inability to predict income level" (Gerhart & Newman, 2020, p. 320). From the agency theory perspective (Jensen & Meckling, 1976), performance-contingent compensation schemes transfer risk from the principal (i.e., firm, organization) to the agent (i.e., employee). In contrast, fixed pay involves

minimal risk such that earnings to employees are guaranteed regardless of performance as long as employment continues. As such, Milgrom and Roberts (1992) described risk averse individuals to be those who "would rather have a smaller income whose magnitude is certain than an uncertain income that is somewhat larger on average but is subject to unpredictable and uncontrollable variability" (p. 187). Therefore, risk averse participants might have sorted into safer (or less risky) pay conditions compared to their relatively high creative performance, which are the three cases of above the diagonal in **Table 11**. On the other hand, risk tolerant participants might have sorted into more risky pay conditions than their actual creative performance, which are the three cases of below the diagonal in **Table 11**.

To examine the influence of risk aversion on sorting decisions, I asked participants to make ten paired lottery decisions on an instrument developed by Holt and Laury (2002)[18] to measure their *risk aversion.* It was asked after completing all four task sessions and debriefing questions to prevent participants from being primed to be more/less risk averse and make biased decisions (Cadsby et al., 2007). As **Table 12** presents, each of the paired decisions consists of Option A, which is relatively safer but provides a smaller expected pay amount, and Option B, which is relatively riskier but provides a larger expected pay amount. Participants chose either Option A or Option B and will have a switching point where they move from a safer option to a riskier option. The higher the value of the switching point, the more risk averse a participant is. In addition to the earnings from the task sessions, participants also received additional payoffs according to their lottery decisions. I found that 36 participants who chose less risky pay conditions (i.e., above the diagonal) showed significantly higher risk aversion (M = 6.72, SD =

[18] Payoffs in the current study are adjusted for inflation using the U.S. Bureau of Labor Statistics consumer price index (CPI) for all urban consumers from the original payoffs in Holt and Laury (2002).

1.91) than the rest of the participants by 16% (M = 5.66, SD = 2.20) and 78 participants who sorted into more risky pay conditions by 21% (i.e., below the diagonal; M = 5.28, SD = 2.04).

Discussion

Until recently, there was a debate about whether PFP enhances or undermines performance in intrinsically interesting tasks. Kim et al.'s (2022) meta-analysis addressed this, showing that PFP has a positive effect on performance (broadly defined) in both noninteresting and interesting tasks and that the magnitudes were similar. However, as they and others have pointed out, the relationship between PFP and creative performance remains unresolved (e.g., Amabile, 1998; Deci & Ryan, 1985; Hendijani, Bischak, Arvai, & Dugar, 2016; Park et al., in press; Pink, 2011; Zhou & Hoever, 2014). Thus, they called for future research that directly examines this question.

Importantly, given that we already know that different types of performance respond differently to PFP as Kim et al. (2022) found smaller effect sizes of PFP on performance quality than performance quantity and as the componential model of creativity suggested (Amabile & Pratt, 2016), we cannot presume that the effect of PFP on creative performance may operate the same direction or be of same magnitude as that on other task performance. Furthermore, the extant research stream on PFP and performance in general that uses a piece-rate scheme (a form of individual incentive rewarded typically based on objective performance evaluation and objective payoff formula) is limited in providing implications on the PFP-creative performance relationship due to the following reasons. First, whereas a piece-rate scheme typically uses objective performance evaluation for physical or easily quantifiable outputs, creative (and other more workplace common) performance relies significantly instead on subjective evaluation of performance. Second, not only how performance is measured, but also how such performance is

tied to payoffs is largely different for creative work settings. That is, while a piece-rate scheme typically adopts objective and absolute payoff formula, in creative or more common work settings, how the payoff is linked to a performance rating is often subjective and relative. Third, because of the reasons I have stated, we cannot simply generalize the sorting mechanism observed in non-creative task settings using a piece-rate to the PFP-creative performance relationship. Similarly, we have not known how and to what extent task autonomy moderates this relationship under the settings closer to typical workplaces.

My study aims to respond to the call by Kim et al. (2022) and others to address and help resolve the competing theoretical arguments on how PFP influences creative performance. In so doing, although my study was conducted in a laboratory setting, I designed the experiment to better reflect the above mentioned key workplace dimensions, such as using creative tasks more similar to those in the workplace, subjective performance evaluation, relative performance rating, and PFP benchmarked from many U.S. companies. Also of major importance, my design, provided opportunities for subjects to move (sort) between pay conditions. My design also moved beyond the simple dichotomy of presence versus absence of PFP to a more refined approach examining how people react differently to varying strengths of PFP (i.e., high PFP intensity, low PFP intensity, and fixed pay).

The results supported Hypothesis 1b that PFP increased creative performance (by 29%) compared to fixed pay. Further, high PFP intensity more strongly enhanced creative performance (by 34%) than did low PFP intensity (by 23%), consistent with traditional motivation and revised creativity theory in support of Hypothesis 2b. Additionally, supplementary analyses revealed that the effects of PFP on creative performance do not depend on which creative task of the two was manipulated. Thus, the current study provides greater generalizability of findings in terms of the

task features by demonstrating that PFP has a consistent, positive effect on creative performance across the two different creative tasks aimed to involve different dimensions of work capabilities (e.g., web search, writing, imagination, divergent thinking).

Next, in Hypothesis 3, I predicted task autonomy as a key moderator to resolve the theoretical debate over the PFP-creative performance relationship. Both recent interpretations of SDT (Gagné & Deci, 2005), as well as revised creativity theory (Amabile & Pratt, 2016) suggest that PFP can enhance (not harm) creative performance when PFP is perceived to be less controlling (i.e., when experienced task autonomy is higher) of one's motivation and behavior. Supporting this, I found a positive moderating effect of task autonomy on the PFP-creative performance relationship such that the positive effect of creative performance under PFP was more strongly positive as task autonomy increased. However, contrary to these theories concerns about harmful effects of PFP under low task autonomy, the positive effect of PFP, although weaker, was always positive here as well. This is perhaps consistent with the findings of the Kim et al. (2022) meta-analysis that the effect of PFP on performance (defined more broadly, mostly in terms of non-creative tasks) is generally positive and rarely negative. A final observation on the role of autonomy is that my results seem to indicate that increasing task autonomy alone, without using PFP in conjunction, may pose the greater risk of not achieving higher creative performance.

I also incorporated the sorting effect of PFP, which has been ignored in in the intrinsic motivation and creativity literatures and studied primarily using piece-rate systems (rare in the workplace) in the broader financial incentives literature. This is an important gap because even if there exist detrimental effects of PFP, as SDT typically argues (sometimes excepting where there is autonomy), the sorting mechanism is expected to help reduce or even remove them (Fang &

Gerhart, 2012), to the degree that those who prefer to work under PFP are able to sort into PFP jobs. To better reflect real labor market dynamics where there is tremendous mobility between jobs, suggesting that workers can to a significant degree sort in and out of jobs based on their preferences, my laboratory experiment included a within-subjects factor that allowed participants to self-select into fixed pay, low PFP intensity, or high PFP intensity.

Consistent with my Hypotheses 4 and 5, I found that actual and self-assessed creative performance each significantly influenced one's sorting decisions. Specifically, the higher actual and self-assessed creative performance, the more likely the choice of higher PFP intensity. Further, supporting Hypothesis 6, I found that self-assessed creative performance partially mediated the relationship between actual creative performance and pay choice, thus helping better understand what explains sorting behaviors. Finally, in Hypothesis 7, I predicted that the effect of PFP on creative performance would be greater in the pay self-selection session, where both the incentive and sorting effects of PFP operated simultaneously to enhance creative performance, than in the random assignment sessions, where only the incentive effect of PFP operated. I found that, once the number of observations were accounted for, when the opportunity for mobility was allowed, the effect of PFP was approximately 3 (23% to 66% in comparison of low PFP with fixed pay) to 16 times (29% to 459% in comparison of PFP with fixed pay) greater than when it was not. Further, decomposition of the sorting and incentive effects from the total effect of PFP revealed that 65% to 94% of the total effect of PFP was attributed to the sorting effect. These results demonstrate that any study that ignores the sorting effect may miss a considerable (indeed the largest) portion of the positive effect of PFP on creative performance.

Theoretical Implications

There has been disagreement over the effect of PFP on creative performance among major motivation and creativity theories. On the one hand, traditional motivation theories posit that intrinsic and extrinsic motivation can be combined additively. Thus, it is predicted that PFP (i.e., extrinsic motivator) enhances performance in intrinsically motivating, creativity requiring tasks (i.e., intrinsic motivator). On the other hand, SDT claims that PFP undermines intrinsic motivation and thus creative performance because PFP is perceived to control one's self-determination and autonomous motivation (e.g., Deci, 1972; Deci et al., 2017; Deci & Ryan, 1985). Early work on creativity theory took a consistent view with the SDT perspective such that extrinsic motivators harm intrinsic motivation that leads to better creative outcomes. (e.g., Amabile, 1998; Lepper et al., 1973). However, revised creativity theory acknowledged that PFP can "boost" intrinsic motivation and creative performance under specific circumstances, such as when PFP is implemented in an autonomy-supportive setting (Hennessey & Amabile, 2010).

Although two meta-analyses were conducted to test the competing theories, due to their study design and scope, the debate still remains unresolved (Byron & Khazanchi, 2012; Patall et al., 2008). My study aims to remedy this, contributing to compensation, motivation, and creativity literature. For instance, I examined creative performance as a dependent variable, whereas Patall et al. (2008) focused on intrinsic motivation as a dependent variable. Per the componential model of creativity theory (Amabile, 1983a; Amabile & Pratt, 2016) and the findings from previous studies (George, 2007; Grant & Berry, 2011), intrinsic motivation is not equivalent to creativity. Thus, a study that explicitly measures creative performance is needed. Moreover, the two previous meta-analyses defined independent variables very broadly, embracing all forms of rewards. Indeed, the most common form of reward was non-contingent (on creative performance) and non-monetary rewards (e.g., verbal praise and extra course credit).

Further, they included primary studies conducted with children subjects in elementary or pre-schools who carried out non-work-like, simple tasks. In contrast, the current study focused solely on PFP designed to closely resemble how actual U.S. companies pay their workers. Additionally, participants were adults and they performed more complex, workplace-like creative tasks.

This is a notable advancement from previous laboratory experiments where designs were less likely to capture key properties of workplaces. It is crucial to make a theoretical distinction between work and non-work settings because people engage in each situation with different initial motives and objectives. That is, people may join in non-work activities such as education and volunteering not necessarily to earn an income, whereas in their jobs, a primary focus must typically include sustaining their standard of living by earning an adequate monetary return (and often with a focus also on an equitable return) from their work in addition to intrinsic motivation/fulfillment (Adams, 1963; Leana & Meuris, 2015; Maslow, 1943). In this regard, my study provides an experimental setting and measures closer to workplace contexts, which I hope can more directly answer the question of how people react to PFP in performing creative work and provide more accurate direction and effect size of this relationship.

I manipulated varying strengths of PFP to better understand the contradicting theoretical predictions. This is because, when PFP intensity is higher, according to SDT and traditional creativity theory, the undermining effect of PFP on creative performance is expected to be exacerbated, whereas, based on traditional motivation and revised creativity theories, the positive effect of PFP on performance is predicted to be greater under higher PFP intensity. Consistent with the latter, I found that creative performance was enhanced even more as PFP intensity increased. As such, the results were consistent with the predictions from traditional motivation theories and revised creativity theory, but not supporting those from SDT. Furthermore, by

showing that high and low PFP intensity have differing effects on creative performance, my study highlights the importance of taking a more refined approach in fully understanding the PFP-performance, creative performance particularly, relationship.

Next, I examined task autonomy as a key moderator that can perhaps not only resolve the contradicting theoretical predictions but also help us understand why the effect of PFP on creative performance varies from study to study. Both the evolving perspectives on creative theory and SDT propose that extrinsic motivators can have a synergistic effect on intrinsic motivation and thus enhance creative performance under an autonomous work environment (Amabile & Pratt, 2016; Gagné & Deci, 2005). The results revealed that, under high PFP intensity relative to fixed pay, creative performance was higher by 47% under high task autonomy, whereas the creative performance increase (22%) due to high PFP intensity was still substantial, but just under one-half as large under low task autonomy. Hence, even under the same high PFP intensity and working on the same creative task, we may expect a substantial difference in the effect of PFP on creative performance, depending upon whether the work is done under high versus low task autonomy. As such, by examining different task autonomy conditions and varying strengths of PFP intensity simultaneously, my study provides a more comprehensive understanding of the effectiveness of PFP on creative performance.

The current study's within-subjects design where participants carry out multiple task sessions allows examining not only the varying incentive effects of different PFP intensities but also the sorting effect of PFP on creative performance. Like in the labor market, especially in cases where workers have mobility opportunities, participants in my experiment were able to determine their preferred pay conditions and paid accordingly. This moves beyond previous studies that solely considered the incentive effect of PFP by randomly assigning a pay condition

to subjects and not allowing them to choose how they were paid. The sorting mechanism suggests another way to resolve the theoretical debate over the effects of PFP such that it may alleviate (if not completely eliminate) the concern over any possible detrimental effect of PFP on motivation and creative performance.

My study unpacks how the sorting process operates, by investigating the interplay between self-assessed creative performance and actual creative performance to influence sorting decisions. Specifically, I found that self-assessed creative performance mediated the positive relationship between actual creative performance and the higher likelihood of sorting into PFP. In turn, the results revealed that the sorting effect accounted for a substantial portion (65% to 94%) to the (total) effect of PFP on creative performance. Thus, any studies that neglect the sorting effect of PFP significantly underestimate the positive effect that PFP may have on creative performance that could be observed when workers are mobile as in the actual labor market. Finally, my study is the first to demonstrate such large sorting effects in a setting where greater uncertainty exists, due to the nature of creative tasks, subjective performance evaluation, no formula-based payoff to performance, and the use of relative performance ratings. However, these design features are much closer to the workplace than is the case in most previous studies that used non-creative tasks with a piece-rate scheme typically characterized as objective and easily quantifiable performance measures, absolute performance ratings, and an objective and straightforward payoff formula. In sum, my study advances compensation, motivation, and creativity literatures by examining how these competing theories' predictions using a design that is closer to and captures more key elements of the workplace.

Practical Implications

There have been mixed views on how money affects employee creativity among not only researchers but also practitioners. On the one hand, Daniel Pink argued in one of his best-selling books, *Drive*, that incentives "can extinguish intrinsic motivation" and "can crush creativity" (p. 87). Similarly, Jens-Uwe Meyer claimed in his TED talk (2019) that "money is the worst thing that can happen to innovators. Money chokes or even kills creativity … Protect innovators from receiving too much money!" On the other hand, we observe that most innovative companies have very competitive compensation schemes in terms of both the overall pay level and PFP. For instance, reports from *Glassdoor* (glassdoor.com) as of January 2022 show that software engineers at Google make on average $150,737 and a maximum of $271,075 per year. That is, employees in creativity requiring jobs working in one of the most innovative firms earn much higher than average U.S. workers ($61,044)[19]. Furthermore, Laszlo Bock, the former Senior Vice President of People Operations at Google, implied a strong PFP scheme at Google, in his book, *Work Rules!* He said, "At Google, we do have situations where two people doing the same work can have a hundred times difference in their impact, and in their rewards. For example, there have been situations where one person received a stock award of $10,000, and another working in the same area received $1,000,000" (p. 205).

Given that employee creativity is believed to be a key resource for organizational innovation, growth, and success, contradicting perspectives on how creative work should be compensated become a pressing concern to be resolved for management. Furthermore, it has been argued that creativity is increasingly needed not only in those jobs that typically require

[19] May 2020 National Occupational Employment and Wage Estimates in the United States showed that the annual average wage among all occupations was $56,310, which is $61,044 in November 2021 after adjusting for CPI.

creativity (e.g., software engineers) but also “at all levels and different types of jobs, including those that may not have traditionally required employees to be creative” (Shalley et al., 2009, p. 489). In this regard, the current study provides important insights into how contemporary organizations should design their compensation plans to enhance employee creativity.

These insights are as follows. First, by documenting the positive effect of stronger PFP intensity on creative performance, my study suggests that organizations may wish to consider whether their current PFP intensity is sufficiently strong. Second, even if concerns remain about a possible detrimental effect of PFP, my study indicates that any such effect is likely to be weakened or even eliminated by the sorting mechanism of PFP. As also noted above, my study showed that the sorting effect accounts for a substantial portion of the total effect of PFP on creative performance. Hence, organizations should not underestimate the predicted positive effect of PFP by considering the creative performance increase (e.g., by only observing the effect of PFP on current workers, ignoring how PFP selectively attracts and retains workers). In other words, for instance, a Google employee who is less capable and thus earns a $10,000 value stock option, compared to $1,000,000, is more likely to move to another firm that has weaker PFP schemes, whereas the availability of a stock option worth $1,000,000 is likely to retain the high performer and attract other high performers. In the absence of such PFP intensity, sorting effects would be quite different and result in a workforce better matched to fixed and equal pay, independent of performance.

Finally, my study suggests that organizations should consider implementing other organizational practices that can complement the effect of PFP on creative performance. In particular, my study suggests that organizations should contemplate building an autonomy-enhancing work environment to boost the positive effects of PFP on creative performance. This

is because the current study showed that, even when PFP intensity was high, creative performance under low task autonomy was not improved as much as it could have been under high task autonomy. Hence, even if organizations implement strong PFP schemes with the intention to enhance employee creativity, they might not be able to enjoy the full advantage from such compensation practices if they do not provide sufficient work autonomy to their employees. For example, organizations can increase autonomy by giving workers more freedom to decide what projects to participate in, how to carry out projects, and/or how to allocate resources without putting too much pressure on them (Amabile, 2016).

Limitation and Future Research

Although the use of a controlled laboratory experiment provides important advantages, it is important to acknowledge the differences between the experiment setting and that of the workplace and thus the challenges for generalizability. First, the pay at stake and the duration of the employment should be considered. The absolute amount of payouts to the participants in this study was commensurate with (or higher than) the local hourly pay rate. However, the money at stake will not be the same as what workers earn biweekly, monthly, or annually. This could weaken the effects of overall pay level or PFP on motivation and thus creative performance. Similarly, with the shorter duration of the experiment than in the case of an actual job, there is no career promotion effect, which may also weaken the performance-enhancing effects of PFP one would find in the workplace, especially considering that promotion provides greater salary growth by moving to an upper pay grade (Trevor et al., 1997).

Therefore, future research should consider use of a longer timeframe or consider career advancement. One option could be using field data that contains changes in pay schemes, employee turnover, promotion, and creative performance. Of course, internal validity (inferring

cause-effect) would be more challenging (Shadish et al., 2002) in the absence of random assignment and control/manipulation of key causal variables (e.g., PFP). However, both types of research, with their strengths and limitations will be needed. As another consideration, given that Jenkins et al.'s (1998) meta-analysis on the relationship between incentives and performance showed almost double the stronger effect of incentives in field settings ($\hat{\rho}$ = .48) than experiment settings ($\hat{\rho}$ = .24), the effect sizes of PFP found in the current experiment could be conservative in terms of what might be found in the field. One additional way to come closer to a field setting in a laboratory setting would be for researchers to inform participants they will only be re-invited to future sessions if they are among the top performers.

Second, participants in the laboratory experiment did not face any labor market frictions in sorting into different pay conditions, whereas workers in the workplace do to varying degrees encounter searching and switching costs when they move between jobs. Thus, it remains to be determined whether and to what degree the essentially frictionless sorting in the experimental setting may overstate the sorting effect in the actual labor market where frictions do exist. However, as noted, more than half of the U.S. workers changed their jobs in 2020 (57% of the annual total separation rate, JOLTS). Further, over the course of a career, the opportunities for sorting increase. Similarly, information asymmetry in understanding various pay schemes did not exist in the experiment, whereas it is likely that job candidates and even current employees often have less than perfect knowledge of a pay scheme in real work settings. In sum, future research that examines the sorting mechanism of PFP and how it influences creative performance in workplace settings would be useful. Specifically, how do mobility costs and imperfect information influence the magnitude of the sorting effects of PFP on creative (and other) performance?

Third, SDT argues that perceived self-competence through (positive) feedback is another key factor to enhance one's self-determination and intrinsic motivation (e.g., Deci et al., 2017), in addition to the perceived autonomy. Creativity theory also posits that PFP can increase intrinsic motivation and creative performance when it appropriately enhances employees' knowledge of their own ability (Amabile & Pratt, 2016). However, due to the relatively short duration of the laboratory experiment and the nature of the tasks that require substantial time for the judges to evaluate creative performance, the current study design could not provide performance feedback to participants between the task sessions. Thus, in this study, the participants had to make sorting decisions knowing neither their creative performance nor their peers'. It is possible that the matching between creative performance and pay conditions would have been stronger if they had received timely feedback. In turn, this might increase the magnitude of the positive effect of PFP on creative performance via assisting participants in making sorting decisions that match their creative performance with pay schemes better. Thus, again, this may translate into the current study providing more conservative estimates of the sorting effects of PFP than what could be found in the setting where timely feedback is given.

Fourth, although the positive incentive and sorting effects of PFP on creative performance can be large and, as in my study, larger with stronger PFP, there are potential design and execution potential pitfalls that can and do arise in workplace settings in using PFP (Gerhart, 2023). For example, for positive sorting effects to occur when stronger PFP is implemented, there must be worker mobility (Gerhart & Feng, 2022). In its absence, dissatisfied employees will remain rather than moving to a better matching organization. As another example, the stronger the PFP intensity, the more important it is that performance measurement

and payoffs are perceived by employees as credible and fair, especially to the degree work is interdependent (Trevor et al., 2012).

Conclusion

Researchers and practitioners alike have put forth competing arguments regarding the effect of PFP on creative performance. To help resolve this debate and provide a more comprehensive understanding of how and when PFP enhances or undermines creative performance, my study investigated varying strengths of PFP, moving beyond the dichotomy of PFP versus fixed pay, examined task autonomy as a key moderator, and incorporated the sorting mechanism, which plays a major role in the labor market but has been ignored in the intrinsic motivation and creativity literatures. Using a laboratory experiment that seeks to better capture (relative to previous research on PFP and creative performance) key aspects of the workplace, my study showed that the stronger the PFP intensity, the greater the positive effect of PFP on creative performance. The results also highlighted the importance of task autonomy as a moderator, such that the positive effect of PFP on creative performance was substantially greater under higher task autonomy, relative to under low task autonomy. Finally, I found that actual creative performance influenced sorting decisions through self-assessed creative performance. As expected, when sorting was permitted, it resulted in a much larger positive effect of PFP on creative performance than when such sorting was not permitted. As such, past research using designs that do not incorporate sorting are likely to underestimate the total effect of PFP. In sum, my study provides important evidence on how PFP influences creative performance. This evidence indicates that PFP enhances creative performance, and this positive effect is larger when task autonomy is high. As noted, these findings have important implications for both theory and practice.

References

Adams, J. S. 1963. Toward an understanding of inequity. ***Journal of Abnormal and Social Psychology***, 67: 422–436.

Alvesson, M. 2004. ***Knowledge-work and knowledge-intensive firms***. New York, NY: Oxford University Press.

Amabile, T. M. 1982. Social psychology of creativity: A consensual assessment technique. ***Personality Processes and Individual Differences***, 43: 997–1013.

Amabile, T. M. 1983a. The social psychology of creativity: A componential conceptualization. ***Journal of Personality and Social Psychology***, 45: 357–376.

Amabile, T. M. 1983b. ***The social psychology of creativity***. New York, NY: Springer-Verlag.

Amabile, T. M. 1993. Motivational synergy: Toward new conceptualizations of intrinsic and extrinsic motivation in the workplace. ***Human Resource Management Review***, 3: 185–201.

Amabile, T. M. 1998. How to kill creativity. ***Harvard Business Review***, 76: 76–87.

Amabile, T. M. 2016. ***KEYS to creativity and innovation***. Center for Creative Leadership.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. 1996. Assessing the work environment for creativity. ***Academy of Management Journal***, 39: 1154–1184.

Amabile, T. M., Hill, K. G., Hennessey, B. A., & Tighe, E. M. 1994. The work preference inventory: Assessing intrinsic and extrinsic motivational orientations. ***Journal of Personality and Social Psychology***, 66: 950–967.

Amabile, T. M., & Pratt, M. G. 2016. The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. ***Research in Organizational Behavior***, 36: 157–183.

Anad, N., Gardner, H. K., & Morris, T. 2007. Knowledge-based innovation: Emergence and embedding of new practice areas in management consulting firms. ***Academy of Management Journal***, 50: 406–428.

Appelbaum, E., Bailey, T., Berg, P., & Kalleberg, A. 2000. ***Manufacturing advantage: Why high performance work systems pay off.*** New York, NY: Cornell University Press.

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. 2009. Large stakes and big mistakes. ***Review of Economic Studies***, 76: 451–469.

Baer, J. 2010. Is creativity domain specific? ***The Cambridge handbook of creativity*** (J. C. Kaufman&R. J. Sternberg): 321–341. New York, NY: Cambridge University Press.

Baer, M., Oldham, G. R., & Cummings, A. 2003. Rewarding creativity: When does it really matter? ***Leadership Quarterly***, 14: 569–586.

Bandura, A. 1997. ***Self-efficacy: The exercise of control***. New York, NY: W. H. Freeman.

Barnes, C. M., Hollenbeck, J. R., Jundt, D. K., DeRue, D. S., & Harmon, S. J. 2011. Mixing individual incentives and group incentives: Best of both worlds or social dilemma? ***Journal of Management***, 37: 1611–1635.

Batey, M., Furnham, A., & Safiullina, X. 2010. Intelligence, general knowledge and personality as predictors of creativity. ***Learning and Individual Differences***, 20: 532–535.

Binnewies, C., Ohly, S., & Sonnentag, S. 2007. Taking personal initivative and communicating about ideas: What is important for the creative process and for idea creativity? ***European Journal of Work and Organizational Psychology***, 16: 432–455.

Blumberg, M., & Pringle, C. D. 1982. The missing opportunity in organizational research: Some implications for a theory of work performance. ***Academy of Management Review,*** 7: 560–569.

Boggiano, A. K., & Ruble, D. N. 1979. Competence and the overjustification effect: A developmental study. ***Journal of Personality and Social Psychology***, 37: 1462–1468.

Boiché, J., & Stephan, Y. 2014. Motivational profiles and achivement: A prospective study testing potential mediators. ***Motivation and Emotion***, 38: 79–92.

Bretz, R. D., Ash, R. A., & Dreher, G. F. 1989. Do people make the place? An examination of the attraction-selection-attrition hypothesis. ***Personnel Psychology***, 42: 561–581.

Brown, K. A., & Huber, V. L. 1992. Lowering floors and raising ceilings: A longitudinal assessment of the effects of an earnings-at-risk plan on pay satisfaction. ***Personnel Psychology***, 45: 279–311.

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

Cable, D. M., & Judge, T. A. 1994. Pay preferences and job search decisions: A person-organization fit perspective. ***Personnel Psychology***, 47: 317–348.

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

Campbell, D. J. 1984. The effect of goal-contingent payment on the performance of a complex task. ***Personnel Psychology***, 37: 23–40.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. 2014. Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. ***Psychological Bulletin***, 140: 980–1008.

Davis, S. J., & Haltiwanger, J. 1999. Gross job flows. In O. Ashenfelter & D. Card (Eds.), ***Handbook of labor economics***, 2711–2805. New York, NY: Elsevier.

deCharms, R. 1968. ***Personal causation***. New York, NY: Academic Press.

Deci, E. L. 1972. The effects of contingent and noncontingent rewards and controls on intrinsic motivation. ***Organizational Behavior and Human Performance***, 8: 217–229.

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***, 125: 627–668.

Deci, E. L., Olafsen, A. H., & Ryan, R. M. 2017. Self-determination theory in work organizations: The state of a science. ***Annual Review of Organizational Psychology and Organizational Behavior***, 4: 19–43.

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

Dodd, N. G., & Ganster, D. C. 1996. The interactive effects of variety, autonomy, and feedback on attitudes and performance. ***Journal of Organizational Behavior***, 17: 329–347.

Dohmen, T., & Falk, A. 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. ***American Economic Review***, 101: 556–590.

Eisenberger, R., & Armeli, S. 1997. Can salient reward increase creative performance without reducing intrinsic creative interest? ***Journal of Personality and Social Psychology***, 72: 652–663.

Eisenberger, R., Pierce, W. D., & Cameron, J. 1999. Effects of reward on intrinsic motivation-Negative, neutral, and positive: Comment on Deci, Koestner, and Ryan (1999). ***Psychological Bulletin***, 125: 677–691.

Eisenberger, R., & Rhoades, L. 2001. Incremental effects of reward on creativity. ***Journal of Personality and Social Psychology***, 81: 728–741.

Eisenberger, R., Rhoades, L., & Cameron, J. 1999. Does pay for performance increase or decrease perceived self-determination and intrinsic motivation? ***Journal of Personality and Social Psychology***, 77: 1026–1040.

Erat, S., & Gneezy, U. 2016. Incentives for creativity. ***Experimental Economics***, 19: 269–280.

Fang, M., & Gerhart, B. 2012. Does pay for performance diminish intrinsic interest? ***International Journal of Human Resource Management***, 23: 1176–1196.

Friedman, R. S. 2009. Reinvestigating the effects of promised reward on creativity. ***Creativity Research Journal***, 21: 258–264.

Furnham, A., Zhang, J., & Chamorro-Premuzic, T. 2005. The relationship between psychometric and self-estimated intelligence, creativity, personality and academic achievement. ***Imagination, Cognition and Personality***, 25: 119–145.

Gagné, M., & Deci, E. L. 2005. Self-determination theory and work motivation. ***Journal of Organizational Behavior***, 26: 331–362.

Garden, A-M. 1991. Relationship between burnout and performance. ***Psychological Reports,*** 68: 963–977.

George, J. M. 2007. Creativity in organizations. ***Academy of Management Annals***, 1: 439–477.

Gerhart, B. 1999. Human resource management and firm performance: Measurement issues and their effect on causal and policy inferences. In P. M. Wright, L. D. Dyer, J. W. Boudreau, & G. T. Milkovich (Eds.), ***Research in personnel and human resources management, Supplement 4***: 31–51. Greenwich, CT: JAI Press.

Gerhart, B. 2007. Horizontal and vertical fit in human resource systems. In C. Ostroff & T. A. Judge (Eds.), ***Perspectives on organizational fit***: 317–346. New York, NY: Lawrence Erlbaum Associates Publishers.

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

Gerhart, B. 2023. ***Compensation*** (14th ed.). New York: McGraw-Hill.

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. 1990. Organizational differences in managerial compensation and financial performance. ***Academy of Management Journal***, 33: 663–691.

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

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

Gerhart, B., Wright, P. M., McMahan, G. C., & Snell, S. A. 2000. Measurement error in research on human resources and firm performance: How much error is there and how does it influence effect size estimates? ***Personnel Psychology,*** 53: 803–834.

Gibbons, R., & Murphy, K. J. 1999. Relative performance evaluation for chief executive officers. ***Industrial and Labor Relations Review,*** 43: 30–52.

Glover, J. A., & Zimmer, J. W. 1982. Procedures to influence levels of questions asked by students. ***Journal of General Psychology***, 107: 267–276.

Gneezy, U., & Rustichini, A. 2000. Pay enough or don't pay at all. ***Quarterly Journal of Economics***, 115: 791–810.

Grant, A. M., & Berry, J. W. 2011. The necessity of others is the mother of invention: Intrinsic and prosocial motivations, perspective taking, and creativity. ***Academy of Management Journal***, 54: 73–96.

Grohman, M., Wodniecka, Z., & Klusak, M. 2006. Divergent thinking and evaluation skills: Do they always go together? ***Journal of Creative Behavior***, 40: 125–145.

Guilford, J. P. 1967. ***The nature of human intelligence***. New York: Cambridge University Press.

Gupta, N., & Shaw, J. D. 1998. Let the evidence speak: Financial incentives are effective!! ***Compensation & Benefits Review***, 30: 26–32.

Hackman, J. R., & Oldham, G. R. 1976. Motivation through the design of work: Test of a theory. ***Organizational Behavior and Human Performance***, 16: 250–279.

Hendijani, R., Bischak, D. P., Arvai, J., & Dugar, S. 2016. Intrinsic motivation, external reward, and their effect on overall motivation and performance. ***Human Performance***, 29: 251–274.

Hennessey, B. A. 1989. The effects of extrinsic constraints on children's creativity while using a computer. ***Creativity Research Journal***, 2: 151–168.

Hennessey, B. A., & Amabile, T. M. 2010. Creativity. ***Annual Review of Psychology***, 61: 569–598.

Holmstrom, B. 1979. Moral hazard in teams. ***Bell Journal of Economics,*** 13: 324–340.

Holt, C., & Laury, S. 2002. Risk aversion and incentive effects. ***American Economic Review***, 92: 1644–1655.

Howard, J. L., Gagné, M., & Bureau, J. S. 2017. Testing a continuum structure of self-determined motivation: A meta-analysis. ***Psychological Bulletin***, 143: 1346–1377.

Ilgen, D. R. 1986. Laboratory research: A question of when, not if. In E. A. Locke (Eds.), ***Generalizing from laboratory to field settings: Research findings from industrial-organizational psychology, organizational behavior, and human resource management***. Lexington, MA: Lexington Books.

Jaussi, K. S., & Dionne, S. D. 2003. Leading for creativity: The role of unconventional leader behavior. ***The Leadership Quarterly***, 14: 475–498.

Jenkins, G. D., 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.

Jovanovic, B., & Moffitt, R. 1990. An estimate of a sectoral model of labor mobility. ***Journal of Political Economy***, 87: 972–990.

Kachelmeier, S. J., Reichert, B. E., & Williamson, M. G. 2008. Measuring and motivating quantity, creativity, or both. ***Journal of Accounting Research***, 46: 341–373.

Kim, J. H., Gerhart, B., & Fang, M. 2022. Do financial incentives help or harm performance in interesting tasks? ***Journal of Applied Psychology***, 107: 153–167.

Kohn, A. 1993. Why incentive plans cannot work. ***Harvard Business Review***, 71: 54–63.

Latham, G. P. 2007. ***Work motivation: History, theory, research, and practice*** (2nd ed.). Thousand Oaks, CA: Sage.

Lazear, E. P. 2000. Performance pay and productivity. ***American Economic Review***, 90: 1346–1361.

Lazear, E. P., & Rosen, S. 1981. Rank-order tournaments as optimum incentive labor contracts. ***Journal of Political Economy,*** 89: 841–864.

Lazear, E. P., & Spletzer, J. R. 2012. Hiring, churn, and the business cycle. ***American Economic Review***, 102: 575–579.

Leana, C., & Meuris, J. 2015. Living to work and working to live: Income as a driver of organizational behavior. ***The Academy of Management Annals***, 9: 55–95.

Lepper, M. R., Greene, D., & Nisbett, R. E. 1973. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. ***Journal of Personality and Social Psychology***, 28: 129–137.

Malik, M. A. R., Butt, A. N., & Choi, J. N. 2015. Rewards and employee creative performance: Moderating effects of creative self-efficacy, reward importance, and locus of control. ***Journal of Organizational Behavior,*** 36: 59–74.

Maslow, A. H. 1943. A theory of human motivation. ***Psychological Review***, 50: 370–396.

Meyer, J.-U. 2019. ***How money kills creativity***. TEDxUniversity of Nicosia.

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

Mumford, M. D., & Gustafson, S. B. 1988. Creativity syndrome: Integration, application, and innovation. ***Psychological Bulletin***, 103: 27–43.

Park, N. K., Chun, M. Y., & Lee, J. 2016. Revisiting individual creativity assessment: Triangulation in subjective and objective assessment methods. ***Creativity Research Journal,*** 28: 1–10.

Park, T-.Y., Eaglesham, R., Shaw, J. D., & Burton, M .D. in press. Beyond productivity: Incentive effects on alternative outcomes. ***Research in Personnel and Human Resource Management.***

Patall, E. A., Cooper, H., & Robinson, J. C. 2008. The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. ***Psychological Bulletin***, 134: 270–300.

Pink, D. H. 2011. ***Drive: The surprising truth about what motivates us***. New York, NY: Riverhead Books.

Podsakoff, P. M., & Organ, D. W. 1986. Self-reports in organizational research: Problems and prospects. ***Journal of Management,*** 12: 531–544.

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

Pretz, J. E., & McCollum, V. A. 2014. Self-perceptions of creativity do not always reflect actual creative performance. ***Psychology of Aesthetics, Creativity, and the Arts***, 8: 227–236.

Runco, M. A., & Acar, S. 2012. Divergent thinking as an indicator of creative potential. ***Creativity Research Journal***, 24: 66–75.

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

Ryan, R. M., & Deci, E. L. 2000b. Intrinsic and extrinsic motivations: Classic definitions and new directions. ***Contemporary Educational Psychology***, 25: 54–67.

Ryan, R. M., & Deci, E. L. 2017. ***Self-determination theory: Basic psychological needs in motivation, development, and wellness***. New York, NY: Guilford Press.

Rynes, S. L., Schwab, D. P., & Heneman III, H. G. 1983. The role of pay and market pay variability in job application decisions. ***Organizational Behavior and Human Performance***, 31: 353–364.

Schaefer, M. 2018. To dramatically increase creativity, slash your budget in half. ***Mark Schaefer: Rise above the noise***. https://businessesgrow.com/2018/01/30/increase-creativity/.

Schneider, B. 1987. The people make the place. ***Personnel Psychology***, 40: 437–453.

Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. ***Experimental and quasi-experimental designs for generalized causal inference***. Boston, MA: Houghton, Mifflin and Company.

Shalley, C. E. 1995. Effects of coaction, expected evaluation, and goal setting on creativity and productivity. ***Academy of Management Journal***, 38: 483–503.

Shalley, C. E., Gilson, L. L., & Blum, T. C. 2009. Interactive effects of growth need strength, work context, and job complexity on self-reported creative performance. ***Academy of Management Journal***, 52: 489–505.

Shalley, C. E., Zhou, J., & Oldham, G. R. 2004. The effects of personal and contextual characteristics on creativity: Where should we go from here? ***Journal of Management***, 30: 933–958.

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

Shaw, J. D., Dineen, B. R., Fang, R., & Vellella, R. F. 2009. Employee-organization exchange relationships, HRM practices, and quit rates of good and poor performers. ***Academy of Management Journal***, 52: 1016–1033.

Skinner, B. F. 1953. ***Science and human behavior***. New York, NY: MacMillan.

Sternberg, R. J. 1985. Implicit theories of intelligence, creativity, and wisdom. ***Journal of Personality and Social Psychology***, 49: 607–627.

Sue-Chan, C., & Hempel, P. S. 2015. The creativity-performance relationship: How rewarding creativity moderates the expression of creativity. ***Human Resource Management,*** 55: 637–653.

Trevor, C. O., Gerhart, B., & Boudreau, J. W. 1997. Voluntary turnover and job performance: Curvilinearity and the moderating influences of salary growth and promotions. ***Journal of Applied Psychology***, 55: 44–61.

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.

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

Weibel, A., Rost, K., & Osterloh, M. 2010. Pay for performance in the public sector—Benefits and (hidden) costs. ***Journal of Public Administration Research and Theory***, 20: 387–412.

WorldatWork. 2018. ***2018 Incentive pay practices for publicly-traded companies***. Deloitte.

Wright, P. M., Gardner, T. M., Moynihan, L. M., Park, H. J., Gerhart, B., & Delery, J. E. 2001. Measurement error in research on human resources and firm performance: Additional data and suggestions for future research. ***Personnel Psychology,*** 54: 875–901.

Yoon, H J., Sung, S. Y., & Choi, J. N. 2015. Mechanisms underlying creative performance: Employee perceptions of intrinsic and extrinsic rewards for creativity. ***Social Behavior and Personality,*** 43: 1161–1180.

Zhang, Y., He, W., Long, L., & Zhang, J. 2022. Does pay for individual performance truly undermine employee creativity? The different moderating roles of vertical and horizontal collectivist orientations. ***Human Resource Management,*** 61: 21–38.

Zhang, Y., Long, L., & Zhang, J. 2015. Pay for performance and employee creativity: The importance of procedural justice and willingness to take risks. ***Management Decision,*** 53: 1378–1397.

Zhou, J. 1998. Feedback valence, feedback style, task autonomy, and achievement orientation: Interactive effects on creative performance. ***Journal of Applied Psychology***, 83: 261–276.

Zhou, J., & Hoever, I. J. 2014. Research on workplace creativity: A review and redirection. ***Annual Review of Organizational Psychology and Organizational Behavior***, 1: 333–359.

Zhou, J., & Oldham, G. R. 2001. Enhancing creative performance: Effects of expected developmental assessment strategies and creative personality. ***Journal of Creative Behavior***, 35: 151–167.

Table 1
Three Pay Conditions

1A. Fixed pay

Performance evaluation rating	Base pay	Incentive	Total payment (Base pay + Incentive)	Expected payment
Top 30%	$3.90	$0	$3.90	$3.90
Next 50%				
Bottom 20%				

1B. Low PFP intensity

Performance evaluation rating	Base pay	Incentive	Total payment (Base pay + Incentive)	Expected payment
Top 30%	$3.20	$2.00	$5.20	$4.30
Next 50%		$1.00	$4.20	
Bottom 20%		$0	$3.20	

1C. High PFP intensity

Performance evaluation rating	Base pay	Incentive	Total payment (Base pay + Incentive)	Expected payment
Top 30%	$2.50	$5.20	$7.70	$5.11
Next 50%		$2.10	$4.60	
Bottom 20%		$0	$2.50	

Table 2

Descriptive Statistics (Pilot Study)

	Variable	Obs.	Mean	SD	1	2	3	4	5	6	7	8	9
1	PFP (1 = PFP; 0 = fixed pay)	64	.50	.50									
2	Creative Performance (Slogan task)	16	2.34	.91	.32								
3	Creative Performance (Article task)	16	2.09	.90	.47	. [a.]							
4	Ordinary performance (Inventory task)	12	37.42	20.90	.61*	-.46	.94						
5	Ordinary performance (Reference task)	20	2.95	2.01	.17	-.51	.53	. [a.]					
6	Task need for creativity	32	3.19	1.55	.30	.23	.10	-.48	-.12				
7	Task interestingness	32	3.03	1.28	.23	-.41	-.29	-.65*	.21	.55*			
8	Task enjoyment	32	2.66	1.41	.14	-.10	-.39	. [b.]	.17	.46*	.79*		
9	Task satisfaction	32	2.91	1.23	.28	.18	.44	-.23	.52*	.52*	.60*	.54*	
10	Task challengingness	32	3.41	1.24	-.29	.38	-.20	-.18	-.15	-.02	-.05	.10	-.14

Note.

N = 16. * $p < .05$

[a.] No observation.

[b.] All participants rated the inventory task to be 1 (Not enjoyable or fun at all).

Table 3

Manipulation Check (Pilot Study)

3A. Task Characteristics (*"how do you perceive each task?"* in scale of 1 through 5)

	Noninteresting tasks (N = 16)		Creative tasks											
			Slogan task (N = 7)				Article task (N = 9)				Two creative tasks combined (N = 16)			
	Mean	SD	Mean	SD	t [a.]	r_{pb}^2 [b.]	Mean	SD	t [a.]	r_{pb}^2 [b.]	Mean	SD	t [a.]	r_{pb}^2 [b.]
Need for creativity	1.88	1.02	4.75	.46	11.23***	.95	4.25	.46	5.46***	.81	4.50	.52	10.25***	.88
Interesting	2.44	1.36	3.63	.92	4.25**	.72	3.63	.92	5.02**	.78	3.63	.89	2.82*	.35
Enjoyable/fun	2.25	1.48	3.25	1.04	3.55**	.64	2.88	1.46	2.20†	.41	3.06	1.24	2.09†	.23
Satisfying	2.32	1.25	3.50	.93	2.00†	.36	3.50	.93	2.11†	.39	3.50	.89	2.97**	.37
Challenging	3.50	1.10	3.63	1.30	.81	.09	3.00	1.51	-1.16	.16	3.31	1.40	-.46	.01

3B. Task Autonomy (*"how much control did you have over what work to do?"* in scale of 1 through 5)

High autonomy (N = 7)		Low autonomy (N = 9)			
Mean	SD	Mean	SD	t [c.]	r_{pb}^2 [b.]
3.71	.76	2.22	1.09	3.07**	.40

3C. Pay Condition (*"how much your pay was dependent on how you perform?"* in scale of 1 through 3)

Fixed pay (N = 32)		PFP (N = 32)			
Mean	SD	Mean	SD	t [d.]	r_{pb}^2 [b.]
1.56	.72	2.50	.67	5.40***	.32

Note.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

[a.] Comparison of noninteresting tasks with creative tasks.

[b.] Squared point-biserial correlation; Proportion of variance explained.

[c.] Comparison of high task autonomy with low task autonomy.

[d.] Comparison of fixed pay with PFP.

Table 4
Descriptive Statistics (Main Experiment)

	Variable	Obs.	Mean	SD	1	2	3	4	5	6	7	8	9
1	Creative performance	872	3.09	1.15									
2	Self-assessed creative performance	872	53.49	27.91	.19*								
3	Fixed pay	266	--	--	-.33*	-.19*							
4	Low PFP intensity	276	--	--	.03	.00	-.45*						
5	High PFP intensity	330	--	--	.28*	.19*	-.52*	-.53*					
6	Task autonomy (1 = high; 0 = low)	218	.50	.50	.15*	.00	.00	.01	-.01				
7	Age	218	22.87	5.10	.07*	.05	.00	-.01	.02	-.03			
8	Male	53	--	--	-.08*	-.06	.01	.00	-.01	.02	-.02		
9	Female	161	--	--	.07*	.04	-.01	.01	.01	-.06	.04	-.95*	
10	Prefer not to identify gender	4	--	--	.05	.06	.00	-.02	.02	.14*	-.08*	-.08*	-.23*

Note.

N = 218.

* $p < .05$

Table 5

Manipulation Check (Main Experiment)

Table 5A. Task Characteristics (*"how do you perceive the task"* in scale of 1 through 5)

	Overall (N = 218)		Slogan (N = 120)		Article (N = 98)			
	Mean	SD	Mean	SD	Mean	SD	t [a.]	r_{pb}^2 [b.]
Need for creativity	4.17	.84	4.28	.79	4.04	.88	2.12*	.41
Interesting	3.61	.99	3.47	1.04	3.78	.90	2.32*	.19
Enjoyable and fun	3.28	1.10	3.23	1.12	3.33	1.07	.67	.10
Satisfying	3.55	.81	3.46	.83	3.65	.79	1.72†	.22
Challenging	3.25	1.11	3.44	1.11	3.02	1.06	2.84*	.02

Table 5B. Task Autonomy (*"how much control did you have over what work to do?"* in scale of 1 through 5)

High autonomy (N = 108)		Low autonomy (N = 110)			
Mean	SD	Mean	SD	t [c.]	r_{pb}^2 [b.]
3.98	.79	2.29	.90	14.75***	.50

Table 5C. Pay Condition (*"how much your pay was dependent on how you perform?"* in scale of 1 through 3)

Fixed pay (N = 218)		Low PFP intensity (N = 218)				High PFP intensity (N = 218)					
Mean	SD	Mean	SD	t [d.]	r_{pb}^2 [b.]	Mean	SD	t [e.]	r_{pb}^2 [b.]	t [f.]	r_{pb}^2 [b.]
1.31	.53	2.22	.61	19.52***	.64	2.73	.50	28.33***	.79	10.54***	.34

Note.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

[a.] Comparison of slogan task with article task.

[b.] Squared point-biserial correlation; Proportion of variance explained.

[c.] Comparison of high task autonomy with low task autonomy.

[d.] Comparison of fixed pay with low PFP intensity.

[e.] Comparison of fixed pay with high PFP intensity.

[f.] Comparison of low PFP intensity with high PFP intensity.

Table 6

Repeated-measures ANOVA of Pay Conditions on Creative Performance

Pay condition	Mean (SD)	In comparison with	
		Fixed	High PFP
Low PFP	3.24 (1.01)	*Model 1*: $F(1, 217) = 55.26$*** ω^{2} [b.] = .20	*Model 2*: $F(1, 217) = 13.16$*** ω^{2} [b.] = .05
High PFP	3.53 (1.12)	*Model 3*: $F(1, 217) = 108.18$*** ω^{2} [b.] = .33	
PFP [a.]	3.38 (1.08)	*Model 4*: $F(1, 435) = 107.82$*** ω^{2} [b.] = .20	
Fixed	2.63 (1.09)		

Note.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

[a.] High and low PFP intensity combined.

[b.] Omega squared; Proportion of variance explained.

Table 7

Split-plot ANOVA of Interaction between Task Autonomy and PFP on Creative Performance

Table 7A. PFP (High and Low PFP Intensity Combined) vs. Fixed Pay

	SS	*df*	*F*	ω^2 [a.]
Between-subjects: Task autonomy				
Task autonomy	9.47	1	4.99*	.02
Participants | Task autonomy	409.52	216		
Within-subjects: PFP				
PFP	83.33	1	110.65***	.20
Task autonomy * PFP	7.53	1	10.01**	.02
Residual	326.84	434		
Total	844.70	653		

Table 7B. High PFP Intensity vs. Low PFP Intensity vs Fixed Pay

	SS	*df*	*F*	ω^2 [a.]
Between-subjects: Task autonomy				
Task autonomy	17.93	1	9.46**	.04
Participants | Task autonomy	409.52	216		
Within-subjects: PFP intensity				
PFP intensity	92.63	2	64.05***	.22
Task autonomy * PFP intensity	12.78	2	8.84***	.03
Residual	312.41	432		
Total	844.70	653		

Note.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

[a.] Omega squared; Proportion of variance explained.

Table 8

Logistic and Multinomial Logistic Regression Analysis of Creative Performance and Sorting into PFP

Table 8A. Sorting into PFP (High and Low PFP Intensity Combined) over Fixed Pay

	Model 1	Model 2	Model 3
	PFP	PFP	PFP
Actual creative performance	1.15***		.81**
	(.21)		(.27)
Self-assessed creative performance		.08***	.08***
		(.01)	(.01)
Constant	-1.73**	-2.30***	-4.16***
	(.52)	(.48)	(.85)
χ^2	43.84***	102.02***	112.45***
df	1	1	2
-2 log-likelihood	186.00	127.81	117.38
Pseudo R^2	.19	.44	.49

Table 8B. Sorting into High PFP Intensity or Low PFP Intensity over Fixed Pay

	Model 1		Model 2		Model 3	
	Low PFP	High PFP	Low PFP	High PFP	Low PFP	High PFP
Actual creative performance	.85***	1.37***			.72**	1.02**
	(.23)	(.23)			(.27)	(.31)
	[2.33]	[.3.95]			[2.06]	[2.78]
Self-assessed creative performance			.06***	.12***	.06***	.11***
			(.01)	(.02)	(.01)	(.02)
			[1.06]	[1.13]	[1.06]	[1.12
Constant	-1.86**	-2.86***	-2.02***	-5.27***	-3.70***	-7.63***
	(.57)	(.59)	(.48)	(.76)	(.86)	(1.13)
	[.16]	[.06]	[.13]	[.01]	[.02]	[.0005]
χ^2	56.12***		153.64***		166.28***	
df	2		2		4	
-2 log-likelihood	391.93		294.41		281.77	
Pseudo R^2	.13		.34		.37	

Note.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

Base category of all models: Fixed pay. Standard errors are in parentheses. Relative risk ratios are in square brackets for Table 8B.

Table 9

Decomposing Sorting Effect of PFP

						Unweighted			Weighted [b.]		
		Fixed pay	Low PFP	High PFP	PFP [a.]	Δ Low PFP / Fixed	Δ High PFP / Fixed	Δ PFP / Fixed	Δ Low PFP / Fixed	Δ High PFP / Fixed	Δ PFP / Fixed
Random assignment *(Sessions 1-3)*	Obs.	218	218	218	436	23%	34%	29%	23%	34%	29%
	Mean	2.63	3.24	3.53	3.38						
Self-selection *(Session 4)*	Obs.	48	58	112	170	37%	68%	58%	66%	292%	459%
	Mean	2.06	2.83	3.46	3.25						
$\frac{Sorting\ effect}{Total\ effect}$ $= \frac{\text{PFP effect under self−selection} - \text{PFP effect under random assignment}}{\text{PFP effect under self−selection}}$ $= \frac{(\text{sorting effect} + \text{incentive effect}) - \text{incentive effect}}{(\text{sorting effect} + \text{incentive effect})}$						38% [c.]	50%	51%	65%	88%	94%

Note.

[a.] High and low PFP intensity combined.

[b.] The number of observations weighted by multiplying the number of observations with the mean.

[c.] Sorting effect accounts for 38% of total effect (= $\frac{37\%\ \text{low PFP effect under self−selection} - 23\%\ \text{low PFP effect under random assignment}}{37\%\ \text{low PFP effect under self−selection}}$)

Table 10

Category of Motivation for Participation, Example Responses, and Creative Performance

Motivation for participation	Example responses	Frequency (%)	Mean (SD)				
			All	High PFP	Low PFP	PFP	Fixed pay
Extrinsically motivated	* For money making, of course * I like winning gift cards * It is a paid study * Interested in the extra cash * I wanted to receive the reward for this study	71 (32.6%)	3.06 (1.01)	3.64 (1.16)	2.92 (.80)	3.25 (1.03)	2.58 (.77)
Intrinsically motivated	* I thought it sounded interesting * I like research * I truly enjoy participating in studies * I thought it would allow me to think outside the box and sounded fun * I was intrigued by the statement that these tasks would be similar to that of an employee	60 (27.5%)	2.96 (1.13)	3.18 (1.17)	3.32 (1.07)	3.27 (1.10)	2.41 (1.00)
Both extrinsically and intrinsically motivated	* Cool study to participate and I can win some money * Sounded fun and also the money doesn't hurt * It sounded very intriguing and the payment seemed very fair * I enjoy participating in studies! I'm also in it for the money * Gift cards are always a great incentive. Also madly interested to see what my brain would come up with	75 (34.4%)	3.18 (1.18)	3.14 (1.17)	3.58 (1.05)	3.38 (1.12)	2.58 (1.19)
Uncategorizable responses	* No particular reason * I received an email * Because it is available to me * I have extra time * Bored out of my mind	12 (5.5%)	3.00 (1.06)	3.04 (1.02)	2.75 (.35)	2.97 (.88)	3.06 (1.51)

Table 11

Examination of Self-assessed Creative Performance and Risk Aversion among Mis-sorted Participants

		Sorting			Self-assessed creative performance (*t*)		Risk aversion (*t*)	
Actual creative performance		High PFP	Low PFP	Fixed pay	Rest of the participants [a.] (N = 182)	Participants below the diagonal [b.] (N = 78)	Rest of the participants [a.] (N = 182)	Participants below the diagonal [b.] (N = 78)
Top 30%	Earned	$7.70	$5.20	$3.90				
	Gain or loss	$0	-$2.50	-$3.80				
	N	51	16	4	-4.14***	-6.00***	2.70**	3.57***
Next 50%	Earned	$4.60	$4.20	$3.90				
	Gain or loss	+$.40	$0	-$.30				
	N	49	25	16				
Bottom 20%	Earned	$2.50	$3.20	$3.90				
	Gain or loss	-$1.40	-$.70	$0				
	N	12	17	28				

Note.

The gray-shaded diagonal cells are the base to calculate the gain and loss from mis-sorting. For example, 16 of top 30% performers earned only $5.20, losing $2.50, by sorting into low PFP intensity who could have earned $7.70 if they sorted into high PFP intensity.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

[a.] Comparison of the three cells above the gray-shaded diagonal with rest of the participants.

[b.] Comparison of the three cells above the gray-shaded diagonal with the three cells below the gray-shaded diagonal.

Table 12

Risk Aversion Measure: Paired Lottery Choice Decisions

	Option A				Option B			
Pair No.	Probability	Payoff	Probability	Payoff	Probability	Payoff	Probability	Payoff
1	10%	$4.20	90%	$3.40	10%	$8.10	90%	$0.20
2	20%	$4.20	80%	$3.40	20%	$8.10	80%	$0.20
3	30%	$4.20	70%	$3.40	30%	$8.10	70%	$0.20
4	40%	$4.20	60%	$3.40	40%	$8.10	60%	$0.20
5	50%	$4.20	50%	$3.40	50%	$8.10	50%	$0.20
6	60%	$4.20	40%	$3.40	60%	$8.10	40%	$0.20
7	70%	$4.20	30%	$3.40	70%	$8.10	30%	$0.20
8	80%	$4.20	20%	$3.40	80%	$8.10	20%	$0.20
9	90%	$4.20	10%	$3.40	90%	$8.10	10%	$0.20
10	100%	$4.20	0%	$3.40	100%	$8.10	0%	$0.20

Note.

Adapted from Holt and Laury (2002).

Figure 1

Comparison between Previous and the Current Laboratory Experiments

Previous laboratory experiments	Current laboratory experiment
Non-adult subjects in non-workplace settings (e.g., preschool, elementary school children)	Adult subjects (aged between 18 and 63, similar to working population)
Non-work-like, simple tasks (e.g., generating titles for a short story that a child imagines being a popcorn, listing novel uses for common physical objects)	Work-like creative tasks (creating advertising slogans, writing magazine articles after conducting web search)
Non-monetary rewards (e.g., extra course credits, verbal praise, candies)	Monetary rewards
Less realistic pay schemes (e.g., piece-rate of $0.01, absence of base pay)	More realistic pay schemes (benchmarked from common incentive rates among U.S. companies; total payout ranged between $13.1 and $28.7 for less than an hour)
Between-subjects design; Incentive effect only	Within-subjects design; Both incentive and sorting effects

Figure 2

Predicted Effect of PFP on Performance, by Theory and Task Type

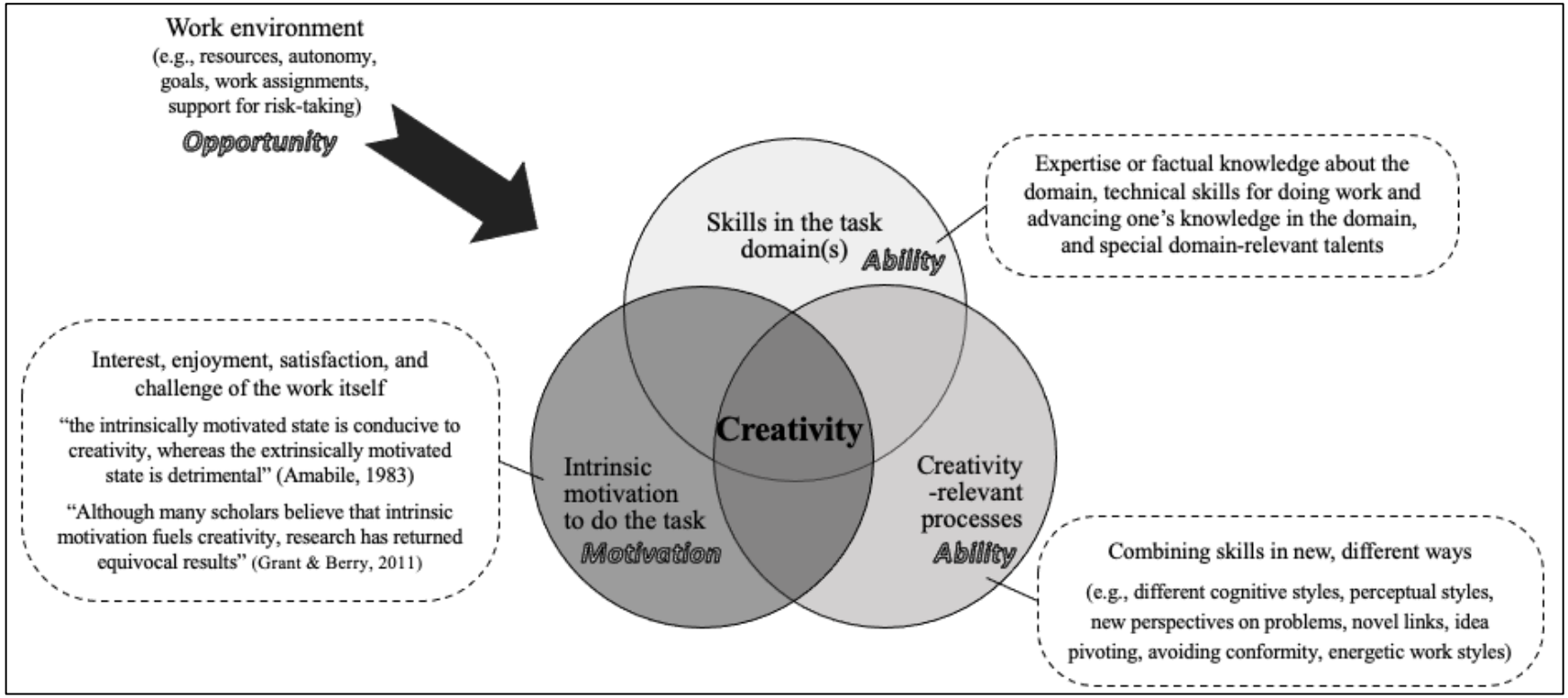
			Theory				
			CET/ SDT	Traditional motivation [a.]	Traditional creativity	Revised creativity	
Task type	Non-creative	Noninteresting	+	+			*Agree on a positive effect of PFP*
		Interesting	-	+			*Disagree on the effect of PFP; Kim et al. (2022) found a positive effect*
	Creative		-	+	-	+	*Not resolved yet; Current study*

Note.

[a.] Expectancy (Vroom, 1964), reinforcement (Skinner, 1953), and equity (Adams, 1963) theories

Figure 3

Componential Model of Creativity



Note.

Adapted from Amabile (1983b) and Amabile and Pratt (2016).

I included ability, motivation, and opportunity labels (Appelbaum et al., 2000; Blumberg & Pringle, 1982; Gerhart, 2007).

Figure 4

Overview of Experimental Design

				Pay *(Within-subjects)*			
				Session 1	**Session 2**	**Session 3**	**Session 4**
Task, Autonomy *(Between-subjects)*	**Condition #1**	Slogan task	High autonomy	Fixed	Low PFP	High PFP	*Choice*
	Condition #2	Slogan task	Low autonomy	Low PFP	High PFP	Fixed	*Choice*
	Condition #3	Article task	High autonomy	High PFP	Fixed	Low PFP	*Choice*
	Condition #4	Article task	Low autonomy	Low PFP	High PFP	Fixed	*Choice*

Note.

Each task session lasted for maximum 10 minutes.

Figure 5

Three Pay Conditions

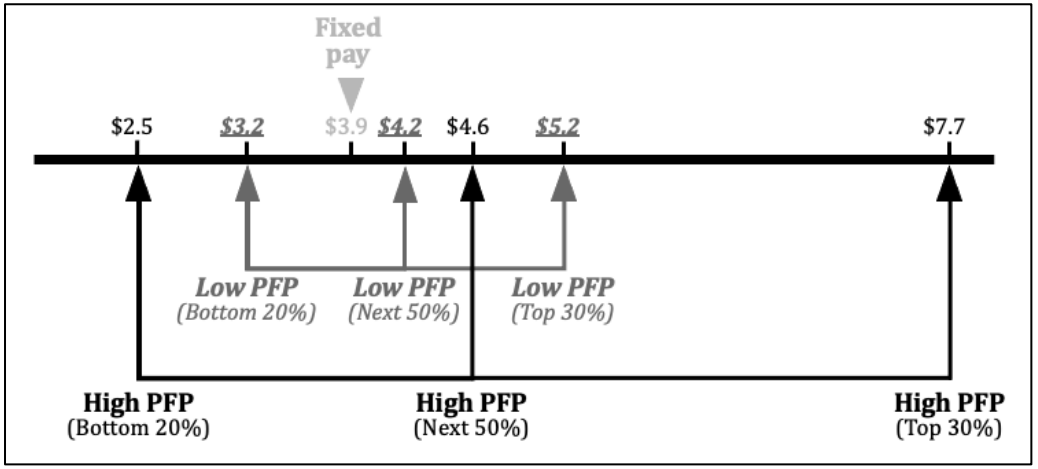


Figure 6

Interaction of Task Autonomy and Pay Conditions on Creative Performance

Figure 6A. PFP (High and Low PFP Intensity Combined) vs. Fixed Pay

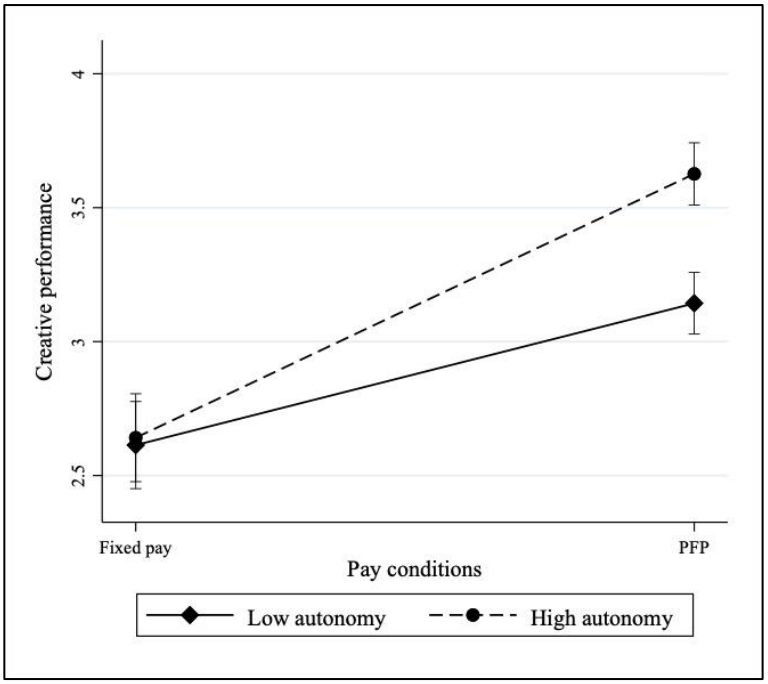


Figure 6B. High PFP Intensity vs. Low PFP Intensity vs. Fixed Pay

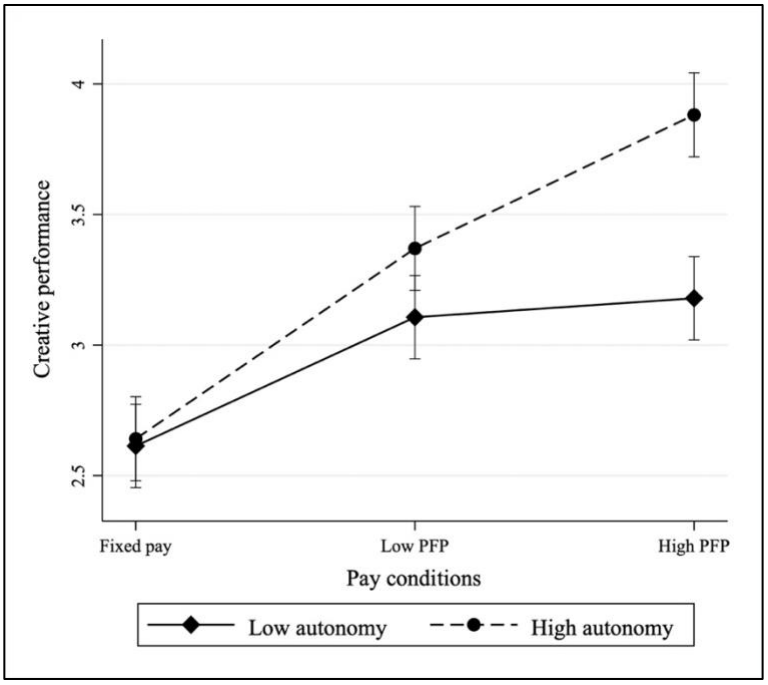


Figure 7

Probability of Sorting into Each Pay Condition Predicted by Actual Creative Performance

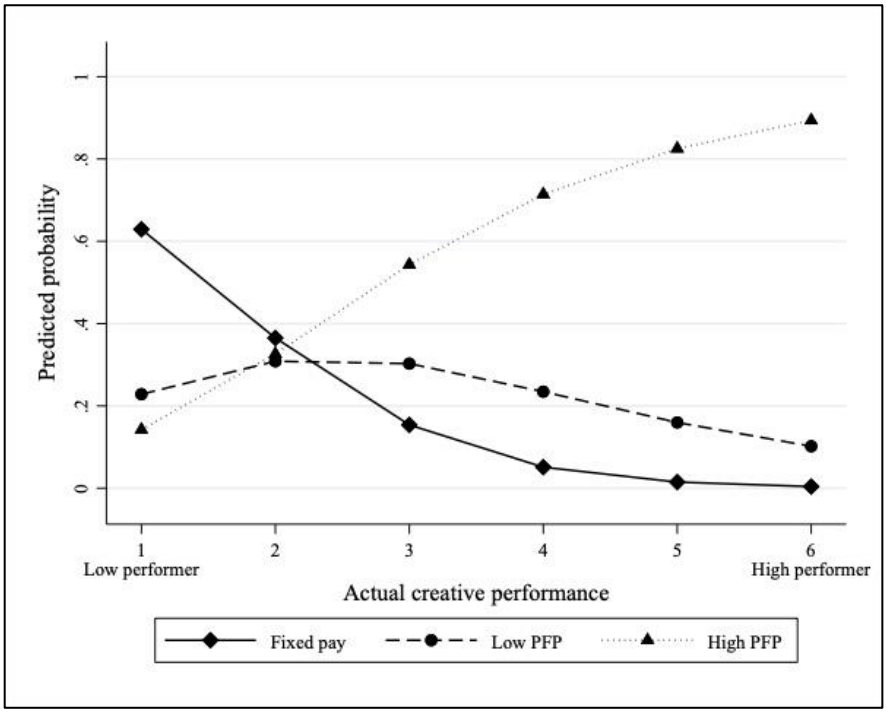
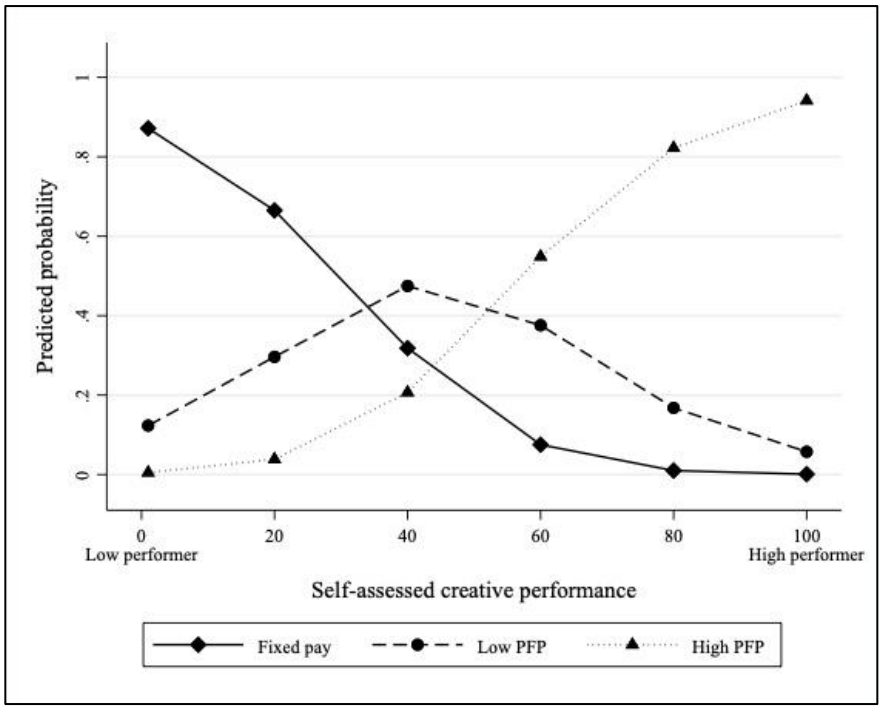


Figure 8

Probability of Sorting into Each Pay Condition Predicted by Self-assessed Creative Performance



Appendix A
An Email to Recruit Participants

Hello.

We are looking for participants for a virtual experiment, which will be held from [Date] to [Date]. You will just need your student ID number and your personal or lab computer to carry out the experiment task. In this study, you will perform a series of tasks that are meant to be similar to those you might perform as an employee in a company. The total time you will spend, including the introduction, 4 experimental sessions, and debriefing questions at the end, will be about 1 hour. You will be compensated with an Amazon Gift Card within 25 days after the completion of the study. The pay (amount of gift card) you will receive depends partially on how you perform in the task, ranging from $13.09 to $28.70 for the entire experiment.

- Sign up at https://uwsob.sona-systems.com/
- Study title: Task Performance Study
- Click on the online study link and you can start when you are ready.

Please feel free to contact me (jkim835@wisc.edu) if you have any questions.

Appendix B

Experimental Instructions for the Slogan Task

Under high autonomy condition:

Three clients have approached you asking for an advertising slogan to help sell their products. You can **freely choose one product** to work on. Choose to work on a slogan for either: (a) a ball launcher for a dog, or (b) a transparent toaster, or (c) a longboard baby stroller.

(a) A ball launcher for a dog	(b) A transparent toaster	(c) A longboard baby stroller
		

Under low autonomy condition (one of three products appears randomly):

A client has approached you asking for an advertising slogan to help sell their product. You **cannot choose** what product to work on. Create a slogan for a ball launcher for a dog.

A ball launcher for a dog

A client has approached you asking for an advertising slogan to help sell their product. You **cannot choose** what product to work on. Create a slogan for a transparent toaster.

A transparent toaster

A client has approached you asking for an advertising slogan to help sell their product. You **cannot choose** what product to work on. Create a slogan for a longboard baby stroller.

A longboard baby stroller

Then, performance evaluation standards are introduced:

You will have 10 minutes to create and submit your advertising slogans. You will see the timer on your screen to check the time elapsed. Please note that the slogan you create will be automatically saved and the experiment window on your computer will automatically take you to next page after 10 minutes and 10 seconds. You will not be able to submit your slogan before the 8 minute mark.

Your performance will be evaluated by the client firm, primarily based on how **creative**, **novel**, and **unique** your slogan is and how **useful** the slogan will be in **increasing product sales volume** and **a premium price** for the product. You may submit a maximum of three advertising slogans, but your performance will be evaluated using the best rated one.

Lastly, the pay condition is presented – high PFP intensity, low PFP intensity, and fixed pay, respectively:

For this particular project, your client will pay you as follows. You will receive $2.50 for participating in this work regardless of your performance. In addition to this base pay, you have a chance to earn an incentive of $5.20 if your slogan is evaluated to be top 30% in its **creativeness** and **usefulness**. In that case, you will receive $2.50 + $5.20 = $7.70. If your work is evaluated to be in the next 50%, the incentive is $1.60, so you would receive $2.50 + $2.10 = $4.60. There is no incentive if your work is evaluated to be bottom 20%. In that case, your total pay = $2.50. The table below summarizes the current compensation plan.

Performance evaluation rating	**Base pay**	**Incentive**	**Total payment (Base pay + Incentive)**
Top 30%	$2.50	$5.20	$7.70
Next 50%		$2.10	$4.60
Bottom 20%		$0	$2.50

For this particular project, your client will pay you as follows. You will receive $3.20 for participating in this work regardless of your performance. In addition to this base pay, you have a chance to earn an incentive of $2.00 if your slogan is evaluated to be top 30% in its **creativeness** and **usefulness**. In that case, you will receive $3.20 + $2.00 = $5.20. If your work is evaluated to be in the next 50%, the incentive is $1.00, so you would receive $3.20 + $1.00 = $4.20. There is no incentive if your work is evaluated to be bottom 20%. In that case, your total pay = $3.20. The table below summarizes the current compensation plan.

Performance evaluation rating	**Base pay**	**Incentive**	**Total payment (Base pay + Incentive)**
Top 30%	$3.20	$2.00	$5.20
Next 50%		$1.00	$4.20
Bottom 20%		$0	$3.20

For this particular project, your client will pay you $3.90 regardless of the **creativeness** and **usefulness** of your slogan. No incentive is available.

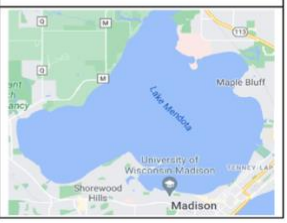
Performance evaluation rating	**Base pay**	**Incentive**	**Total payment (Base pay + Incentive)**
Top 30%	$3.90	$0	$3.90
Next 50%			
Bottom 20%			

Appendix C

Experimental Instructions for the Article Task

Under high autonomy condition:

Editor-in-chief of a local magazine asked you to write a short article on (a) UW-Madison, (b) Bucky Badger, or (c) Lake Mendota. You can **freely choose one topic** to work on. You may search online about interesting and fun facts about the topic and write a short article in a creative way.

(a) UW-Madison	(b) Bucky Badger	(c) Lake Mendota
		

Under low autonomy condition (one of three products appears randomly):

Editor-in-chief of a local magazine asked you to write a short article on UW-Madison. You **cannot choose** what topic to work on. You may search online about interesting and fun facts about the topic and write a short article in a creative way.

UW-Madison

Editor-in-chief of a local magazine asked you to write a short article on Bucky Badger. You **cannot choose** what topic to work on. You may search online about interesting and fun facts about the topic and write a short article in a creative way

Bucky Badger

Editor-in-chief of a local magazine asked you to write a short article on Lake Mendota. You **cannot choose** what topic to work on. You may search online about interesting and fun facts about the topic and write a short article in a creative way

Lake Mendota

Then, performance evaluation standards are introduced:

You will have 10 minutes to search online and submit your article. You will see the timer on your screen to check the time elapsed. Please note that your article will be automatically saved and the experiment window on your computer will automatically take you to next page after 10 minutes and 10 seconds. You will not be able to submit your article before the 8 minute mark.

Your performance will be evaluated by the editor-in-chief, primarily based on how **creative**, **novel**, and **unique** your article is and how **useful** the article will be in **increasing magazine subscriptions** and **the reputation of the magazine**. You may submit a maximum of three articles, but your performance will be evaluated using the best rated one.

Lastly, the pay condition is presented – high PFP intensity, low PFP intensity, and fixed pay, respectively:

For this particular project, the editor-in-chief will pay you as follows. You will receive $2.50 for participating in this work regardless of your performance. In addition to this base pay, you have a chance to earn an incentive of $5.20 if your article is evaluated to be top 30% in its **creativeness** and **usefulness**. In that case, you will receive $2.50 + $5.20 = $7.70. If your work is evaluated to be in the next 50%, the incentive is 2.10, so you would receive $2.50 + $2.10 = $4.60. There is no incentive if your work is evaluated to be bottom 20%. In that case, your total pay = $2.50. The table below summarizes the current compensation plan.

Performance evaluation rating	**Base pay**	**Incentive**	**Total payment (Base pay + Incentive)**
Top 30%	$2.50	$5.20	$7.70
Next 50%		$2.10	$4.60
Bottom 20%		$0	$2.50

For this particular project, the editor-in-chief will pay you as follows. You will receive $3.20 for participating in this work regardless of your performance. In addition to this base pay, you have a chance to earn an incentive of $2.00 if your article is evaluated to be top 30% in its **creativeness** and **usefulness**. In that case, you will receive $3.20 + $2.00 = $5.20. If your work is evaluated to be in the next 50%, the incentive is $1.00, so you would receive $3.20 + $1.00 = $4.20. There is no incentive if your work is evaluated to be bottom 20%. In that case, your total pay = $3.20. The table below summarizes the current compensation plan.

Performance evaluation rating	**Base pay**	**Incentive**	**Total payment (Base pay + Incentive)**
Top 30%	$3.20	$2.00	$5.20
Next 50%		$1.00	$4.20
Bottom 20%		$0	$3.20

The editor-in-chief will pay you $3.90 regardless of the **creativeness** and **usefulness** of your article. No incentive is available.

Performance evaluation rating	**Base pay**	**Incentive**	**Total payment (Base pay + Incentive)**
Top 30%	$3.90	$0	$3.90
Next 50%			
Bottom 20%			

Appendix D
Topic Sets for Each Task Session

D.1. The Slogan Task

(a) A sailing yacht	(b) An air fryer	(c) An AI translator
		

(a) A robot vacuum	(b) A luxury RV	(c) An electric bicycle
		

(a) A ball launcher for a dog	(b) A transparent toaster	(c) A longboard baby stroller
		

(a) A smart ring	(b) A roll up piano keyboard	(c) A massage chair
		

D.2. The Article Task

(a) UW-Madison	(b) Bucky Badger	(c) Lake Mendota
		

(a) Pizza	(b) Hamburger	(c) Hotdog
		

(a) Microwave	(b) Laptop	(c) Vacuum
		

(a) Yellowstone National Park	(b) Grand Canyon National Park	(c) Yosemite National Park
		

Appendix E
PFP Intensity Calculation and Comparison with Previous Experiments

		Fixed pay	Low PFP intensity				High PFP intensity			
		Base pay	Base pay	Incentive	Total payment (Base + Incentive)	Incentive rate	Base pay	Incentive	Total payment (Base + Incentive)	Incentive rate
		$3.90	$3.20				$2.50			
Performance rating	Top 30%			$2.00	$5.20			$5.20	$7.70	
	Next 50%			$1.00	$4.20			$2.10	$4.60	
	Bottom 20%			$0	$3.20			$0	$2.50	
Incentive rate #1 [(Max. total payment in PFP condition - Min. total payment in PFP condition) / Min. total payment in PFP condition]						0.63				2.08
Incentive rate #2 [Max. total earning in PFP condition / total earning in fixed pay condition]						1.33				1.97
Incentive rate #3 [Max. incentive in PFP condition / total earning in fixed pay condition]						0.51				1.33
Incentive rate #4 [Max. incentive in PFP condition/ Max. total earning in PFP condition]						0.38				0.68
Incentive rate #5 [Max. total earning (top performer) in PFP condition / average performer earning in PFP condition]						1.24				1.67

Note.

1. Performance rating cutoffs (i.e., 30% - 50% - 20%) were simplified based on Exhibit 10.5 in Gerhart and Newman (2020).
2. The expected total payment for the entire experiment (i.e., fixed pay, low PFP intensity, high PFP intensity, and a pay choice session) is $17.75, which is very close to Cadsby et al.'s (2007) average payout of $17.28 after adjusting for the currency exchange and CPI. Minimum possible total payment in all four task sessions is $12.10 (= $3.90 under fixed pay + $3.20 under low PFP intensity + $2.50 under high PFP intensity + $2.50 under high PFP intensity if the participant chooses high PFP intensity) which meets the minimum hourly pay policy ($12.00).
3. Because there is no single definitive way to calculate incentive rate to measure PFP intensity, here I lay out five formulae.
4. Incentive rates were benchmarked from a large scale report of compensation among 166 publicly traded U.S. companies (WorldatWork, 2018). The report used the ratio of compensation for average to high performers. Approximately top 20% companies in terms of PFP intensity (i.e., companies with higher PFP intensity) showed the incentive rate between 1.50 and 1.74, meaning that, in these companies, top performers would earn between $15 and $17, and average performers $10. On the other hand, approximately 20% of companies those in the lower end of PFP intensity showed the incentive rate between 1.10 and 1.24, indicating that, in these companies, top performers earn between $11 and $12, average performers $10. Incentive rate #5, which adopted this formula, is set to be 1.67 for high PFP intensity and 1.24 for low PFP intensity.
5. The experimental design in Gneezy and Rustichini (2000) had four pay conditions—fixed pay, low PFP intensity, medium PFP intensity, and high PFP intensity—where subjects were randomly assigned to one of the four pay conditions. One of the most notable differences between the pay design in their study and the one in the current study is that they provided the equivalent base pay (60 NIS: New Israeli Shekel) to every pay condition. This means that the three PFP conditions do not involve any pay at risk compared to the fixed pay condition. However, the base pay in the current study was designed (a) to impose a pay at risk for the PFP conditions by setting lower base pay for high PFP intensity ($2.50), low PFP intensity ($3.20), than fixed pay ($3.90); (b) but to provide a risk premium by awarding sufficient incentives such that total payment for high performers is much higher under high PFP intensity ($7.70), low PFP intensity ($5.20), than fixed pay ($3.90) conditions.
6. In contrast, Ariely et al. (2009) implemented three levels of PFP intensity (low, medium, and high) without fixed pay condition in their first and second experiments. This hinders from examining the effects of PFP on performance compared to those of fixed pay which is claimed to be more desirable pay design by SDT and creativity literature. Moreover, their PFP conditions did not have a base pay which may put too much of pay risk as opposed to the case of Gneezy and Rustichini (2000) that did not have any pay risk. In addition, given that companies usually guarantee, at least, minimum wage as a base pay, not providing any base pay could be a less realistic and too extreme experimental setting.

Appendix F

Experimental Instructions for the Inventory Task

You have a chance to fill an inventory list for a food processing company. Your role is to sort and fill in the inventory list so that distribution team can send products to Stores A, F, G, and K. The first letter of the serial number indicates the store codes where products should be sent to. You should sort a list in ascending order (A to Z) of the store to be sent, the date of production (earlier to later), the date of in-stock (earlier to later), product quality (A- to A++), and weight (lighter to heavier). You will be given 5 sets of lists one by one, and each set has 10 items to be sorted.

You will be given 10 minutes to work on. You will see the timer on your screen to check the time elapsed. Please note that your work will be automatically saved, and the experiment window on the computer will automatically take you to next page after 10 minutes and 10 seconds. You will not be able to submit your work before the 9 minute mark in the session.

Your performance will be evaluated by both the accuracy of the list and the number of items you sorted.

Set (1):

Item #	Serial No.	Date of Production	Date of In-stock	Date of Expiration	Weight	Quality
1	A0712889	JAN-03-2020	JAN-18-2020	MAR-03-2021	5.75lb	A+
2	K2346658	APR-15-2020	APR-29-2020	MAY-15-2021	3.45lb	A++
3	A1852263	JAN-03-2020	JAN-10-2020	MAR-03-2021	6.25lb	A
4	F9567445	MAY-23-2020	JUN-01-2020	JUL-11-2021	5.50lb	A
5	G5228588	FEB-12-2020	FEB-27-2020	APR-12-2021	4.80lb	A+
6	K7754999	APR-13-2020	APR-22-2020	MAY-12-2021	6.15lb	A
7	A1853000	JUL-28-2020	AUG-03-2020	OCT-15-2021	7.48lb	A-
8	F9578895	DEC-01-2020	DEC-11-2020	FEB-07-2022	1.23lb	A++
9	F8363211	OCT-25-2020	NOV-04-2020	JAN-20-2022	2.57lb	A
10	G4487893	JAN-17-2020	JAN-30-2020	MAR-13-2021	2.40lb	A-

Lastly, the pay condition is presented – PFP and fixed pay, respectively:

For this particular task, your manager will pay you as follows. You will receive $2.70 for participating in this work regardless of your performance. In addition to this base pay, you have a chance to earn an incentive of $0.10 for every correctly sorted item. Thus, if you correctly sort all 10 items in 5 sets, you will receive $7.70 (= $2.70 base pay + $5.00 incentive).

For this particular task, your manager will pay you $3.80 regardless of the number of correctly sorted items. No incentive is available.

Appendix G

Experimental Instructions for the Reference Task

You have a chance to write a reference list of academic articles for a publisher. You are a proofreader at a publishing company that covers the subject areas of psychology, economics, and engineering. Your role is to write a reference list for articles following the given citation style of each article's subject field. You will be given total of 12 articles to work on within 10 minutes.

You will see the timer on your screen to check the time elapsed. Please note that your work will be automatically saved, and the experiment window on the computer will automatically take you to next page after 10 minutes and 10 seconds. You will not be able to submit your work before the 9 minute mark in the session.

Psychology articles for this specific journal follow *modified APA style*, economics articles *modified Chicago style*, and electrical engineering articles *modified IEEE style*.

Your performance will be evaluated by both the accuracy of the reference and the number of articles you listed.

Article (1):

Rules for modified Chicago style:

- **Subject area:** Economics
- **Author names:** All authors' names should be included. If there are multiple authors, the first author is written as [last name], [first name], and use comma [,] between authors and use comma [,] and [and] before the last author. Except for the first author, author name should be written as [first name] [last name]. Put period [.] at the end of the author names.
- **Publication year:** Comes after author names. Put period [.] after the year.
- **Article titles:** Capitalize the first word and all major words of the titles and subtitles of articles. Short conjunctions (e.g., and, as, so, if), articles (e.g., a, an, the), and short prepositions (e.g., as, by, for, per, to) are minor words. Words of four letters or more are major words (e.g., with, between, from, much, than). Enclose the article title in quotes [" "]. Mark period [.] at the end.
- **Journal titles:** Capitalize all major words in the journal titles.
- **Volume, issue, and page numbers:** Volume numbers are directly put after the journal title with a space in between. Volume numbers are followed by issue numbers between parentheses [( )]. Place space between volume number and issue number. Place colon [:] after the issue numbers and before the page numbers. Page numbers should be put in full numbers, but certain numbers could be omitted if the pages are from the same range (e.g., use 320-41 if pages are from 320 to 341). Mark period [.] at the end.
- **Example:** Lastname, Firstname, Firstname Lastname, and Firstname M. Lastname. Year. "Title of Article." Title of Journal volume number (issue number): page-page.

Subject area: Economics
Title: Disease and development: The effect of life expectancy on economic growth
Author(s): Daron Acemoglu; Simon Johnson
Journal: Journal of Political Economy
Publication year: 2007
Volume number: 115
Issue number: 6
Page: 925 – 985

Lastly, the pay condition is presented – PFP and fixed pay, respectively:

For this particular task, your manager will pay you as follows. You will receive $2.70 for participating in this work regardless of your performance. In addition to this base pay, you have a chance to earn an incentive of $0.40 for every correctly written reference item. Thus, if you correctly sort all 12 articles, will receive $7.50 (= $2.70 base pay + $4.80 incentive).

For this particular task, your manager will pay you $3.80 regardless of the number of correctly written reference items. No incentive is available.

Appendix H

Random Assignments of Experiment Conditions in the Pilot Study

Participant No.	Autonomy	Session 1		Session 2		Session 3		Session 4	
		Task	Pay	Task	Pay	Task	Pay	Task	Pay
1	High	Reference	Fixed	Article	Fixed	Reference	PFP	Article	PFP
2	Low	Reference	PFP	Reference	Fixed	Article	Fixed	Article	PFP
3	High	Inventory	PFP	Inventory	Fixed	Slogan	Fixed	Slogan	PFP
4	Low	Inventory	Fixed	Inventory	PFP	Article	Fixed	Article	PFP
5	High	Reference	Fixed	Slogan	PFP	Slogan	Fixed	Reference	PFP
6	Low	Inventory	PFP	Inventory	Fixed	Slogan	Fixed	Slogan	PFP
7	High	Inventory	Fixed	Article	Fixed	Article	PFP	Inventory	PFP
8	High	Reference	Fixed	Article	Fixed	Article	PFP	Reference	PFP
9	Low	Reference	Fixed	Reference	PFP	Slogan	Fixed	Slogan	PFP
10	Low	Reference	Fixed	Reference	PFP	Slogan	Fixed	Slogan	PFP
11	Low	Reference	Fixed	Reference	PFP	Article	Fixed	Article	PFP
12	Low	Reference	PFP	Reference	Fixed	Article	Fixed	Article	PFP
13	High	Slogan	Fixed	Reference	Fixed	Reference	PFP	Slogan	PFP
14	Low	Reference	PFP	Reference	Fixed	Slogan	Fixed	Slogan	PFP
15	High	Article	PFP	Inventory	PFP	Inventory	Fixed	Article	Fixed
16	Low	Inventory	PFP	Inventory	Fixed	Slogan	Fixed	Slogan	PFP

Appendix I

Mean and Standard Deviation of Creative Performance by Pay Condition and Task Autonomy

Pay Conditions	Sorting? [a.]	Overall	High task autonomy	Low task autonomy
High PFP	N	3.53 (1.12)	3.88 (1.13)	3.18 (1.00)
Low PFP	N	3.24 (1.01)	3.37 (.97)	3.11 (1.04)
PFP [b.]	N	3.38 (1.08)	3.63 (1.08)	3.14 (1.02)
Fixed	N	2.63 (1.09)	2.64 (1.10)	2.61 (1.09)
High PFP	Y	3.46 (1.13)	3.70 (1.09)	3.25 (1.14)
Low PFP	Y	2.83 (1.00)	3.11 (.87)	2.54 (1.06)
PFP [b.]	Y	3.25 (1.13)	3.49 (1.05)	3.01 (1.16)
Fixed	Y	2.06 (.97)	2.08 (1.10)	2.04 (.84)

Note.

Standard deviations are in parentheses.

Two creative tasks are combined.

[a.] N if random assignment of pay conditions; Y is sorting/ self-selection into pay conditions.

[b.] High and low PFP intensity conditions are combined.

Appendix J

Repeated-measures ANOVA of Pay Conditions on Creative Performance

	Pay condition	Sorting? [a.]		*df*	F	ω^2 [b.]
1	[High PFP & Low PFP] vs. [Fixed]	N	Two tasks	(1, 435)	107.82***	.20
2		N	Slogan task	(1, 239)	45.01***	.15
3		N	Article task	(1, 195)	64.43***	.24
4	[High PFP] vs. [Low PFP] vs. [Fixed]	N	Two tasks	(2, 434)	61.43***	.22
5		N	Slogan task	(2, 238)	24.11***	.16
6		N	Article task	(2, 293)	39.31***	.28
7	[High PFP] vs. [Low PFP]	N	Two tasks	(1, 217)	13.16***	.05
8		N	Slogan task	(1, 119)	2.88†	.02
9		N	Article task	(1, 97)	12.50***	.10
10	[High PFP] vs. [Fixed]	N	Two tasks	(1, 217)	108.18***	.33
11		N	Slogan task	(1, 119)	43.53***	.26
12		N	Article task	(1, 97)	68.01***	.40
13	[Low PFP] vs. [Fixed]	N	Two tasks	(1, 217)	55.26***	.20
14		N	Slogan task	(1, 239)	25.26***	.17
15		N	Article task	(1, 195)	30.17***	.23
16	[High PFP & Low PFP] vs. [Fixed]	Y	Two tasks	(1, 217)	44.08***	.17
17		Y	Slogan task	(1, 118)	3.76†	.02
18		Y	Article task	(1, 96)	60.24***	.38
19	[High PFP] vs. [Low PFP] vs. [Fixed]	Y	Two tasks	(2, 215)	30.09***	.21
20		Y	Slogan task	(2, 117)	4.09*	.05
21		Y	Article task	(2, 97)	38.09***	.43
22	[High PFP] vs. [Low PFP]	Y	Two tasks	(1, 168)	12.91***	.07
23		Y	Slogan task	(1, 100)	4.05*	.03
24		Y	Article task	(1, 66)	9.79**	.11
25	[High PFP] vs. [Fixed]	Y	Two tasks	(1, 158)	56.18***	.26
26		Y	Slogan task	(1, 81)	5.90*	.06
27		Y	Article task	(1, 75)	76.61***	.50
28	[Low PFP] vs. [Fixed]	Y	Two tasks	(1, 104)	16.03***	.12
29		Y	Slogan task	(1, 53)	.77	.00
30		Y	Article task	(1, 49)	18.66***	.23

Note.

† $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

[a.] N if random assignment of pay conditions; Y if sorting/ self-selection into pay conditions.

[b.] Omega squared; Proportion of variance explained.