

October 24, 2009

The Impact of Risk Aversion and Stress on the Incentive Effect of Performance Pay

C. Bram Cadsby

Department of Economics

University of Guelph

Guelph, Ontario, Canada

Fei Song

Ted Rogers School of Business Management

Ryerson University

Toronto, Ontario, Canada

Francis Tapon

Department of Economics

University of Guelph

Guelph, Ontario, Canada

Acknowledgements: We gratefully acknowledge funding from SSHRC grants #410-2007-1380 (Cadsby) and #2006-410-138 (Song). We are greatly indebted to J. Atsu Amegashie for many important and stimulating discussions and suggestions. We are also grateful to James Cox, Mike Hoy, Ling Liu, Alex Maynard, Thanasis Stengos and seminar audiences at the University of Guelph, University of Waterloo, University of Tasmania and University of Sydney for many helpful suggestions and comments.

The Impact of Risk Aversion and Stress on the Incentive Effect of Performance Pay

Abstract

We demonstrate that effectiveness of performance-contingent incentives is inversely related to individual risk-aversion levels through two mechanisms: 1) rational optimizing decisions about the amount of effort to supply when effort is positively correlated with risk exposure and 2) the possibly choke-inducing stress accompanying financial uncertainty. In two laboratory studies using real-effort tasks, we find a significant inverse relationship between productivity improvement under performance pay and risk-aversion levels. Moreover, we show that both mechanisms help explain this result. For about 25% of participants, performance actually deteriorates under performance pay, and the probability of such deterioration increases with risk aversion and stress.

Keywords: risk aversion, performance pay, incentive, stress, choking under pressure, productivity, pay for performance, piece rate, experiment, compensation.

JEL Codes: C91, M52, J33.

The effect of financial incentives on human behavior is a central concern of economics (e.g., Baker et al., 1988; Bénabou and Tirole, 2003; Gibbons, 1998; Kreps, 1997; Lazear, 1986, 2000a, 2000b; Prendergast, 1999). Moreover, the relationship between such incentives and work performance has been an important focus of attention not only in economics, but also in the management and psychology literatures (e.g., Gerhart and Rynes, 2003; Gomez-Mejia and Welbourne, 1988; Rynes et al., 2005; Vroom, 1964). However, the impact of such incentives on behavior is viewed quite differently both within and across these disciplines. Many economists regard monetary incentives as a dominant and effective motivator of human behavior. Consequently, they argue that performance-contingent pay can be an effective incentive device to induce desired performance by mitigating principal-agent problems (e.g., Hart, 1989; Holmstrom, 1979; Jensen and Meckling, 1976; Jensen and Murphy, 1990; Kahn and Sherer, 1990; Kale et al., 2009; Milgrom and Roberts, 1992, pp. 206–247; Seiler, 1984). However, other economists and some psychologists argue, often with supporting experimental evidence, that under some circumstances extrinsic monetary incentives can crowd out intrinsic motivation, thereby adversely affecting motivation and hence performance (e.g., Deci and Ryan, 1985; Gneezy and Rustichini, 2000; see Frey and Jegen, 2001 for a review of this literature). Furthermore, a different strand in the psychology literature suggests that extrinsic incentives that increase the subjective importance of performing well on a task can result in “choking under pressure” (e.g., Baumeister, 1984; see Baumeister and Showers, 1986 for a review of this literature), thus hindering performance even when motivation is maintained or enhanced. Recently, economists have begun examining this issue as well (e.g., Ariely et al., 2009; Dohmen, 2008).

The incentive effect of pay-for-performance has been extensively examined in both the laboratory (e.g., Brase, 2009; Cadsby et al., 2007; Fessler, 2003;) and the field (e.g., Abowd, 1990; Banker et al., 1996; Fehr and Goette, 2007; Gerhart and Milkovich, 1990; Lazear, 2000a, 2000b; Paarsch and Shearer, 1999, 2000, 2009; Shearer, 2004; Stajkovic and Luthans, 2001). Although a considerable number of studies show that monetary incentives are effective at improving performance, others indicate no effect on performance, while still others demonstrate that under some circumstances incentives can hinder performance (Bonner et al., 2000; Camerer and Hogarth, 1999; and Jenkins et al., 1998, all provide comprehensive reviews of this literature). All of these empirical studies are primarily concerned with the *average* effect of incentives on performance for a particular population in a specific context. This literature has focused on how the characteristics

of the work task, the relationships both between principals and agents and among agents, and the particular structure of the monetary incentives may affect the presence and magnitude of an incentive effect. However, to our knowledge, none of the earlier studies has explored whether individual differences among agents may also have an important and predictable impact on the effectiveness of financial incentives at improving task performance.¹ In this paper, we focus on one such factor: differences in attitudes toward financial risk and uncertainty.

Whenever the link between effort and performance has a random component, the payoff from the exertion of effort involves financial uncertainty under performance-contingent pay. Indeed the trade-off between the provision of financial incentives and the transfer of risk from risk-neutral principals to risk-averse agents is a central theme of agency theory (e.g. Milgrom and Roberts, 1992; Prendergast, 1999). This trade-off implies that the optimal level of incentive intensity is lower for higher levels of risk and lower levels of risk tolerance. A lower level of incentive intensity is associated with less effort. However, in the standard linear model of agency theory, where the level of uncertainty is determined exogenously and is independent of the effort levels chosen by agents, the optimal amount of effort under a given level of incentive intensity is determined independently of either the amount of risk faced by agents or their attitudes toward it (e.g., Prendergast, 1999; Sloof and Praag, 2008).

Individual attitudes toward financial risk may nonetheless affect work performance under exogenously determined incentives. First, in contrast to the standard model, the amount of financial uncertainty faced by an agent may be positively correlated with his/her effort. In such circumstances, a more risk-averse agent may rationally choose to exert less effort than a less risk-averse counterpart in order to reduce exposure to risk. Less effort then translates into poorer average performance for the more risk-averse agent.

Second, there is empirical evidence that individual differences play an important role in the propensity of individuals to choke under pressure when faced with an academic test or sports competition (see Baumeister and Showers, 1986, pp. 373–375 for a review). Financial incentives may also enhance the perceived importance of performing well and have been associated with a similar choking phenomenon (Baumeister and Showers, 1986, pp. 368–369; Ariely et al., 2009). Since a higher level of risk aversion implies a greater discomfort with financial uncertainty, it may

¹ Paarsch and Shearer (2007) do examine potential gender differences in the response to piece rates for workers from a tree-planting firm in British Columbia, Canada. Controlling for ability, they find no difference.

also be associated with a greater tendency for performance to be impaired by choking on incentivized tasks.

The purpose of this paper is to use experimental data to examine how an individual's attitude toward risk may influence the effectiveness of financial incentives at improving his/her performance on an assigned task. In doing so, we consider both rational decisions about the amount of effort to exert under the financial uncertainty that inevitably accompanies performance pay and the possibly choke-inducing stress experienced under the resultant pressure to perform well. The next section provides some theoretical background. We then report on the two laboratory studies using real-effort tasks with salient financial incentives undertaken to carry out this investigation. Some conclusions follow.

THEORETICAL BACKGROUND

Under a fixed-pay compensation scheme where no performance-contingent incentives are offered, risk is borne by the firm. In contrast, under a performance-contingent incentive scheme, some of the risk is borne by the employees. In particular, a performance-based incentive scheme involves financial uncertainty, exposing the investment of effort by individual employees to financial risk. The reactions of employees to a change from a risk-free fixed-salary scheme (henceforth *fs*) to a riskier pay-for-performance scheme (henceforth *ppp*) may thus differ depending upon individual attitudes toward risk. This may occur for two reasons. First, under *ppp*, a more risk-averse employee may rationally choose to exert less effort than a less risk-averse employee if effort is positively correlated not only with expected output, but also with risk. Second, a more risk-averse employee may react with greater stress than a less risk-averse employee to the financial uncertainty of *ppp*, and this increased stress may hinder his/her performance response to monetary incentives. We discuss both of these arguments below.

i. Optimal Effort and Risk Aversion when Risk Increases with Effort

To examine the relationship between optimal effort and individual levels of risk aversion under *fs* and *ppp*, we employ a simple agency model with a linear piece rate and both additive and multiplicative uncertainty (e.g. Bushman et al., 2000; Baker and Jorgensen, 2003; Sloof and van Praag, 2008). Consider an agent whose level of output, y , depends stochastically on the amount of effort s /he exerts, a , and two random terms, ε and θ :

$$(1) y = \theta \cdot a + \varepsilon,$$

where $\theta \sim N(\mu_\theta, \sigma_\theta^2)$ and $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. His/her pay, w , is determined by a fixed salary component, s , and a piece-rate component, b :

$$(2) w = s + b \cdot y.$$

The cost of effort is measurable in monetary terms and is assumed to be quadratic:

$$(3) c(a) = k \cdot a^2/2 - f \cdot a,$$

where a is effort, while k and f are constants with $k > 0$ and $f \geq 0$. The f term represents intrinsic motivation², allowing for the possibility that effort exerted on the task may be enjoyable or satisfying up to a certain comfort level after which it becomes costly. This allows for some effort to be exerted under fs .³ Each agent is assumed to possess a Constant Absolute Risk Aversion (CARA) utility function:

$$(4) U = -\exp[-r(w - k \cdot a^2/2 + f \cdot a)],$$

where r is the Arrow-Pratt measure of risk aversion.⁴ Since neither source of randomness in the relationship between effort and output is known prior to the exertion of effort, each agent must base his/her decisions on expected utility:

$$(5) E(U) = -\exp\{-r[s - k \cdot a^2/2 + f \cdot a] - r \cdot b \cdot a \cdot \mu_\theta + [(r \cdot b)^2/2] \cdot [a^2 \cdot \sigma_\theta^2 + \sigma_\varepsilon^2]\}.$$

Defining CE as the certainty equivalent in monetary terms of this expected utility and noting that $U(CE) = E(U)$ by definition, we obtain:

$$(6) CE = s - k \cdot a^2/2 + f \cdot a + b \cdot a \cdot \mu_\theta - [(r \cdot b)^2/2] \cdot [a^2 \cdot \sigma_\theta^2 + \sigma_\varepsilon^2].$$

A risk-averse agent maximizing his/her expected utility or its certainty equivalent under piece-rate *ppf* compensation will then determine his/her optimal effort as:

² Like Frey and Jegen (2001) in their survey on motivation crowding out, we employ the definition of intrinsic motivation introduced by psychologist Deci (1971, p. 105) as follows: “one is said to be intrinsically motivated to perform an activity when one receives no apparent reward except the activity itself.”

³ The standard assumption that effort always involves disutility is a special case of this assumption where $f = 0$. In fact, all of the participants in our experimental studies produced some output under *fs* compensation. This may be due to a positive f in individual cost functions. However, there are alternative explanations such as a feeling of reciprocal obligation to do something in exchange for the fixed salary or a desire to practise for subsequent rounds in which the participant may believe s/he could face *ppf* compensation.

⁴ The $f \cdot a$ term representing intrinsic motivation could be introduced as a direct positive component of the utility function rather than as a negative component of the $c(a)$ function. Such an approach yields the same expression for equation (4) and identical theoretical predictions as the approach taken in the text.

$$(7) a_{pfp}^* = \frac{f + b \cdot \mu_\theta}{k + r \cdot b^2 \cdot \sigma_\theta^2} \quad ^5, ^6$$

Note that r , the Arrow-Pratt measure of risk-aversion, is inversely related to the optimal amount of effort. This relationship depends on the variance of the multiplicative random term, σ_θ^2 . Since the multiplicative random term implies a positive correlation between effort and risk, a more risk-averse person will sacrifice more expected return than a less risk-averse person to mitigate risk. The variance of the additive random term, σ_ε^2 , plays no role in the determination of optimal effort because additive risk is independent of the actions of the agent. Any part of income that is designated as fixed salary, s , and is thus paid regardless of performance, also plays no part in the determination of effort. For an agent who is paid solely by means of a fixed salary, the piece rate, b , is zero. For such an agent, optimal effort is:

$$(8) a_{fs}^* = \frac{f}{k}.$$

Thus, in the fs case, optimal effort is independent of the agent's attitude toward risk because there is no financial risk for the agent. Of course, if $f = 0$, there is no intrinsic motivation. In that case, any effort brings disutility to the agent and hence none will be exerted in the fs case.

The incentive effect of pfp is the difference between performance under pfp and performance under fs . This difference in performance is closely related to the difference in effort exerted under the two pay schemes as specified in equation (1). This difference in effort is:

$$(9) a_{pfp}^* - a_{fs}^* = \frac{f + b \cdot \mu_\theta}{k + r \cdot b^2 \cdot \sigma_\theta^2} - \frac{f}{k}.$$

The most important implication of (9) for this paper is that, *ceteris paribus*, there is an inverse relationship between an individual agent's level of risk aversion and the difference in his/her motivation to exert effort under pfp relative to fs . This inverse relationship holds regardless

⁵ This expression was derived for the case where $f = 0$ by Bushman et al. (2000) and Baker and Jorgensen (2003). However, neither of these papers focused on the relationship between individual levels of risk aversion and the optimal level of effort.

⁶ The first-order condition is $-k \cdot a + f + b \cdot \mu_\theta - r \cdot b^2 \cdot a \cdot \sigma_\theta^2 = 0$. The second order condition is $-k - r \cdot b^2 \cdot \sigma_\theta^2 < 0$ whenever $a^* > 0$. This encompasses all cases in which the agent is risk-averse or risk-neutral as well as those where a preference for risk is not too large, specifically where $-r < k / b^2 \cdot \sigma_\theta^2$. A corner solution involving the exertion of the maximum possible level of effort is possible if an agent is sufficiently risk-loving, i.e. if $-r > k / b^2 \cdot \sigma_\theta^2$.

of whether an agent exhibits intrinsic motivation ($f > 0$) or not ($f = 0$).

Moreover, when $f > 0$ so that there is some intrinsic motivation, effort may either rise or fall when an agent moves from fs to pfp compensation. Whether it rises or falls also depends on the agent's degree of risk aversion. Define $\tilde{r} = k \cdot \mu_\theta / f \cdot b \cdot \sigma_\theta^2 > 0$. If $r < \tilde{r}$, both effort and expected performance will be greater under pfp . This includes all agents who are either risk-neutral or risk-loving, together with those who possess levels of risk aversion below the critical value. However, if $r > \tilde{r}$, both effort and expected performance will be lower under pfp . Moreover, a higher level of multiplicative risk, a higher piece rate, and a higher intrinsic motivation term, f , all reduce \tilde{r} . This enlarges the range of risk-aversion levels resulting in a decline in effort and expected performance under pfp .

Such a decline in effort under performance pay closely resembles the crowding out of intrinsic by extrinsic motivation. However, there is a subtle difference. Crowding out involves a reduction in intrinsic motivation through the introduction of pfp . While such crowding out may be an important phenomenon (e.g., Bénabou and Tirole, 2003; Frey and Jegen, 2001), we have demonstrated that even if intrinsic motivation, f , remains unchanged, effort may nonetheless fall in the presence of multiplicative uncertainty for a different reason. Despite unchanged intrinsic motivation, sufficiently risk-averse agents will choose to exert less effort under pfp than under fs in order to reduce the financial uncertainty associated with the pfp compensation scheme. In contrast to crowding out, this phenomenon should be associated empirically with individuals possessing higher levels of risk aversion. Of course, in the case of no intrinsic motivation ($f = 0$), an agent exerts no effort under fs . Thus, effort cannot fall under pfp in such a case.

ii. Choking under Pressure, Risk Aversion and Performance

Besides making a rational decision to exert less effort resulting in a lower expected level of output, a more risk-averse person working under pfp might well experience considerably more anxiety, stress or pressure than a less risk-averse person. A considerable literature exists concerning the relationship between stress and job performance (see LePine et al., 2005 and Muse et al., 2003, for critical reviews of this literature). Much of the empirical literature suggests that stress, particularly “hindrance stress”, is negatively related to performance.⁷ Furthermore, research on job-related stress has identified performance-contingent pay as one of these stressors (Schuler, 1980). If

⁷ However, Muse et al. (2003) argue that the inverted-U theory, which suggests that small amounts of stress aid performance while larger amounts impede it, has not yet been fairly tested.

such stress impedes performance for a given level of motivation and effort, the incentive effect of *pdf* on performance may be weakened, eliminated or even reversed. As outlined in Baumeister and Showers (1986), such “choking under pressure” may occur for a number of reasons. The payoff uncertainty created by the link between payoff and performance may cause a risk-averse employee to become distracted from the task at hand by thoughts irrelevant to the accomplishment of the task. For example, worry about whether or not one will perform well or the financial implications of performing poorly can seriously impede performance (Baumeister and Showers, 1986, pp. 366). Alternatively, the discomfort created by performance-dependent payoffs can make the risk-averse employee become more self-conscious about each step of the performance process, thereby hindering performance (Baumeister, 1984). Thus, in addition to deciding to exert less effort, a more risk-averse person may have more difficulty transforming his/her effort into output owing to excessive performance stress.

iii. Hypotheses

Both the optimal-effort and choking-under-pressure arguments suggest Hypothesis 1, which is the central focus in the two experimental studies that follow.

H1: The effectiveness of performance-contingent incentives at improving performance is inversely related to individual levels of risk aversion.

Study 1 provides corroborating evidence for Hypothesis 1. However, the experimental design does not permit an examination of the two different proposed mechanisms behind the effect of risk aversion on improvement in performance. Study 2 is designed for that purpose. In study 2, each person participates in a real-effort task for four periods under *fs* and for four periods under *pdf*. The individual variances of performance across the four *pdf* rounds can thus be used as an *ex-post* measure of the financial risk exposure chosen by different individuals under *pdf*.⁸ There is of course no financial risk associated with *fs* regardless of the individual variances across the *fs* rounds. Three further hypotheses are tested using data from that study.

H2: If financial risk is positively correlated with effort: (a) Less risk-averse individuals choose more effort and hence more risk exposure under pdf, leading to an inverse relationship between individual levels of risk-aversion and ex-post individual levels of performance variance across rounds in pdf, but not in fs; (b) Since more effort also implies higher expected output, the relationship in (a) leads to a positive correlation between increases in the ex-post individual level

⁸ Variance in pay is equal to $b^2 \cdot \text{var}(y)$ under *pdf*.

of performance variance and improvements in performance from fs to pfp.

H3: (a) More risk-averse individuals report a greater difference in stress levels between pfp and fs compared to less risk-averse individuals. (b) This difference in stress levels related to risk aversion partially explains the inverse effect of risk aversion on productivity improvement.

H4: A higher level of risk aversion and a greater difference in stress levels are both inversely related to the probability that productivity will be higher under pfp than under fs.

Hypotheses 1, 2, and 3 are illustrated in Figure 1.

STUDY ONE⁹

Participants were recruited at the Queensland University of Science and Technology in Brisbane, Australia by means of both announcements in economics classes and random recruitment in the lounge area of the business school. All 115 participants (71 men and 44 women with an average age of 20.9 years and a standard deviation of 4.51 years) were undergraduates and most but not all were majors in economics or other subjects taught within the business faculty.

A widely used anagram word-creation game (Locke and Latham, 1990; Schweitzer et al., 2004; Vance and Colella, 1990) was employed as the experimental task. Specifically, participants were asked to play one practice and eight experimental three-minute anagram games using prescribed sets of seven letters. The experiment utilized two different compensation schemes, one representing *pfp* and the other *fs*. The *pfp* scheme paid \$0.20 per correct word created. The *fs* scheme paid a fixed amount of \$2.20 for each three-minute anagram game, independent of performance.

Upon arrival, the experimental instructions were read to the participants while they followed along on their own copies. Participants were provided with a prepared workbook containing the anagrams. Each anagram was presented on a separate page of the workbook. Other pages were used to explain which compensation scheme would apply in a subsequent anagram round. Participants were not permitted to look ahead to future pages or to go back to previous pages. They were allowed to tear off one page and look at the next only when instructed to do so by the experimenter.

⁹ Data reported in this study come from a larger study focusing on sorting and incentive effects of performance pay that was published in Cadsby et al. (2007). In the current paper, we focus on the middle four rounds of that eight-round study in order to compare differences in within-person productivity between *pfp* and *fs* rounds.

To ensure anonymity, participants wrote their assigned participant numbers, but not their names, on each page of the workbook immediately prior to beginning work on that page.

In this paper, we focus only on the four middle rounds: rounds 3, 4, 5, and 6.¹⁰ In rounds 3 and 5 participants were paid according to the *fs* scheme, while in rounds 4 and 6 participants were paid according to the *ppp* scheme. In each case, they were informed of the payment scheme immediately prior to the round. After every round, each participant's list of words was collected by the experimenters and taken to another room where the number of correct words was calculated. Participants did not receive feedback on the number of correct words they had created until they were paid at the end of the session.

After participants completed the experimental task, they filled out a questionnaire in which they responded to a number of demographic questions such as age, gender, and native language. Besides collecting demographic data, another purpose of the questionnaire was to elicit risk preferences by asking participants to make ten lottery-choice decisions based on an instrument developed by Holt and Laury (2002). Each of the ten lottery decisions presented to the participants involved a relatively safe choice (option A) versus a relatively risky choice (option B). The probabilities of each lottery outcome are manipulated so that each decision involves progressively higher expected earnings for the risky choice relative to the safe choice. Accordingly, everyone should have a switching point, above which safer choices are selected and below which riskier choices are selected. Following Holt and Laury (2002), we used the number of safe choices as a measure of risk aversion. In addition to being paid for the words they created according to the compensation schemes outlined above, each participant was paid an additional sum based on the outcome of the lottery s/he chose from one of the ten lottery pairs selected at random by drawing a number from an envelope in front of the participants.

We elicited risk preferences after the completion of the experimental task in order to avoid the possibility of biasing the behavioral decisions by priming participants to focus on risk. To mitigate any impact that completing the experimental task might have on risk elicitation, we did not give any feedback on how many correct words were created or how much had been earned until the very end of the experiment after the risk data were collected. Holt and Laury (2002) found that risk preferences were affected by the amount of money at stake. In particular, larger stakes were

¹⁰ In the first and last two-rounds, participants were permitted to choose their preferred compensation scheme. These data are analyzed in Cadsby et al. (2007).

associated with a higher level of risk aversion. We therefore adjusted the stakes used by Holt and Laury (2002) to correspond as closely as possible to the amount at stake in the two rounds of the anagram game affected by each self-selection decision. This involved multiplying Holt and Laury's (2002) lottery numbers by 2.2 to obtain the appropriate amounts in Australian dollars. At the end of the session, players were taken individually to another room, where they were paid privately in cash.

All 115 participants completed the study. The dependent variable is the productivity improvement between the *ppp* and *fs* treatments. It is measured as the within-person productivity difference between average performance in the *ppp* rounds, i.e. rounds 4 and 6, and that of the *fs* rounds, i.e. rounds 3 and 5. The independent variable is a measure of risk aversion based on the number of safe choices by each participant using the Holt-Laury (2002) instrument described above. In addition, we initially included dummy variables for gender and whether the participant's first language was English as controls.¹¹ Table 1 reports means, standard deviations, and correlations of the variables.

An examination of the data reveals the following observations. First, most participants showed an improvement in productivity moving from the *fs* rounds to the *ppp* rounds, suggesting a positive incentive effect of performance pay. On average, participants solved 21.12 problems under *ppp* and 18.86 problems under *fs*. Thus, on average, productivity improved by 2.26 problems and this improvement was significant ($df = 114, p < 0.001$). Note that this was the case even though the mean levels of productivity were slightly higher for the particular anagrams used in the *fs* scheme than for those used in the *ppp* scheme in both a study done by Vance and Colella (1990) conducted with no salient financial incentives ($M_{ppp} = 18.48$ vs. $M_{fs} = 18.74$) and our own pre-test using a piece rate ($M_{ppp} = 24.10$ vs. $M_{fs} = 24.24$). There was however a possible confounding factor because the *fs* rounds (rounds 3 and 5) were run prior to the *ppp* rounds (rounds 4 and 6). Thus, participants may have improved with practice. To remove this confound, we also compared productivity in the earlier *ppp* round 4 ($M_{ppp} = 10.46$) with that in the later *fs* round 5 ($M_{fs} = 9.86$) so that any effect of practice on productivity would work in the opposite direction to the predicted effect of financial incentives on performance. Although the productivity difference of 0.6 words was lower than under the previous comparison, suggesting that participants were indeed improving with practice, it was nonetheless still positive and significant ($t = 2.053, df = 114, p = 0.042$).

¹¹ Out of the 30 participants whose first language was not English, 27 indicated Chinese as their first language.

Second, our participants were quite risk-averse with an average of 6.77 safe choices. As indicated in the Study 1 column of Table 2, 93% exhibited some degree of risk aversion. Of those remaining, 3.5% were risk-neutral, while another 3.5% were risk-loving. These levels of risk aversion were somewhat higher than those found by Holt and Laury (2002) in their lower-stakes setting and roughly comparable to those found in their higher-stakes setting.¹² The correlations presented in Table 2 show that productivity improvement from the *fs* rounds to the *ppf* rounds, was significantly correlated with risk aversion. Neither gender nor first language was correlated with either productivity improvement or risk aversion. Thus, gender and first language were both dropped from the subsequent analysis.

Finally, we tested H1, which predicted that the effectiveness of *ppf* at improving productivity would be inversely related to the level of risk aversion, by regressing the within-person performance difference on individual levels of risk aversion. Following Baker *et al.* (2008), Dickinson (2009), Eriksson *et al.* (2009) Galbiati and Vertova (2008), Heinemann *et al.* (2009) and many others, we used the number of safe choices to represent individual levels of risk aversion.¹³ We found a significant and inverse relationship between productivity improvement and the level of risk aversion ($\beta = -0.223$, $p = 0.017$), corroborating H1. Thus, our results show that the incentive to earn more money through better performance was a less effective motivator for more risk-averse individuals as predicted. In fact, 29 out of 115 participants, accounting for 25.2% of our sample, actually experienced a decline in productivity in the *ppf* rounds.

STUDY TWO

Study 2 was designed to achieve two goals. First, earlier literature has shown that the specific task to be accomplished may have an impact on the presence and the size of the incentive effect of financial incentives (e.g. Camerer and Hogarth, 1999). Thus, we felt it prudent to examine the robustness of our Study-1 results of an inverse relationship between individual levels of risk aversion and the magnitude of performance improvement under *ppf* using a different task. Second, we have proposed two different mechanisms that could underlie the relationship between risk

¹² Recall that we multiplied Holt and Laury's (2002) lower stakes setting by 2.2 to approximate the monetary stakes in two rounds of our anagram game. Hence our stakes were in between their lower and higher stakes settings.

¹³ We also calculated the range of Arrow-Pratt risk-aversion coefficients, r , consistent with the choices made by each participant, based on the assumption of CARA utility. We then took the midpoint of that range for each participant and used that as an alternative measure of risk aversion. All of the results in this paper for both Study 1 and Study 2 were qualitatively identical with only inconsequential differences in p -values when these estimates of r were used instead of the number of safe choices. These results are not reported to conserve space, but are available from the authors upon request.

aversion and the magnitude of the incentive effect under *ppf*. The first is that when effort and risk are positively correlated, more risk-averse persons make a strategic choice to exert less effort in order to reduce financial risk, thereby sacrificing expected return in the process. The second is that a higher level of risk aversion could lead to greater stress and anxiety over the uncertainty inherent in performance-contingent compensation, and that this in turn could result in choking under pressure, thus hindering performance. Study 1 showed that the incentive effect of performance-contingent pay was indeed inversely related to individual levels of risk aversion, thus corroborating H1. However, it could shed no light on the relative importance of the two proposed mechanisms because it made no attempt to elicit stress levels experienced by individual participants under the two compensation treatments. Furthermore, there was no way of measuring effort in Study 1. Thus, it was impossible to determine whether effort and financial risk were positively correlated, as is required by our model for optimal effort to be inversely related to individual levels of risk aversion. Moreover, since there were only two assigned *fs* rounds and two assigned *ppf* rounds, there were not enough degrees of freedom to employ an *ex-post* outcome variance measure as a proxy for the riskiness of individual strategies. In Study 2, we both collect self-reported perceptions of stress under *fs* and *ppf* and assign each participant to four rather than just two rounds under each compensation scheme. We also gather data that can be used as a reasonable proxy for effort. This permits an examination of the relative importance of the two proposed mechanisms and their implications, summarized by Hypotheses 2 and 3, presented at the end of the Theoretical Background section. It also allows us to investigate whether higher levels of risk aversion and stress reduce the probability of an improvement in performance under *ppf* relative to *fs* as predicted by Hypothesis 4.

The design of Study 2 followed as closely as possible that of Study 1. The main differences between the two studies are the following. First, we changed the experimental task from a word-creation anagram task to an arithmetic task of adding up sets of five double-digit integers (e.g., Niederle and Vesterlund, 2007). Specifically, participants were asked to play one practice and eight experimental three-minute rounds using randomly generated sets of five double-digit integers. Second, we added measures of stress into the post-experiment questionnaire. We follow the conventional understanding of stress as the personal reaction of a particular individual to stimuli in his/her environment (e.g., Parker and DeCotiis, 1983). Thus, we used a self-report measure that asked participants to indicate how stressful they found working under the *fs* and *ppf* schemes, using

a Likert scale of 1 (not at all) to 5 (extremely). Third, we switched the order of the *pfj* and *fs* rounds so that the *pfj* rounds occurred before the *fs* rounds. Thus we avoided the possibility of a spurious incentive effect due to participants improving with practice since any such improvement would handicap *pfj* relative to *fs* performance under the new ordering of the treatments. Accordingly, rounds 1, 3, 5, and 7 were *pfj* rounds while rounds 2, 4, 6, and 8 were *fs* rounds.

Participants were recruited at the University of Guelph in Guelph, Ontario by means of email solicitation through the Bachelor of Commerce program listserv. All 60 participants were undergraduates and majors in economics or other business subjects. There were 36 males and 24 females with an average age of 19.27 years and a standard deviation of 1.55 years. All 60 participants completed the study. The primary dependent variable, the productivity improvement occurring under performance pay, was measured as the within-person productivity difference between the *pfj* and *fs* rounds. Individual levels of risk aversion were measured as in Study 1.¹⁴ The variance of performance outcomes was calculated for each individual using the four available rounds of performance data from each of the *fs* and *pfj* treatments. The difference in *pfj* and *fs* performance variances was then calculated for each participant. The difference in reported stress levels experienced under *pfj* and *fs* was also employed as an independent variable. We also collected data on the number of questions attempted in the *pfj* treatment. While recognizing that the level of effort exerted on a real-effort task encompasses such unobservable factors as the degree of concentration and care devoted to it in addition to the observable number of attempts, we used the latter as an imperfect proxy for effort. Lastly, we gathered demographic data on gender and whether or not a participant was born in Canada to use as control variables. Table 3 reports means, standard deviations, and correlations of the variables.

An overview of the data yields the following observations. First, most participants showed an improvement in productivity when moving from the *fs* rounds to the *pfj* rounds. On average, participants solved 3.18 more problems ($M_{pfj} = 30.02$ vs. $M_{fs} = 26.84$) and the within-person difference was significant ($t = 3.68$, $df = 59$, $p < 0.001$). However, 17 out of 60 participants, accounting for 28.4% of the sample, experienced a decline in productivity in the *pfj* rounds, while an additional 4 participants, accounting for 6.7%, neither improved nor suffered a decline in

¹⁴ The numbers used on the Holt-Laury instrument were in Canadian rather than Australian dollars and slightly lower in numerical value than the Australian-dollar numbers used in Study 1. In particular, while the original Holt-Laury numbers were multiplied by 2.2 for the Australian-dollar amounts used in Study 1, they were multiplied by 2 for the Canadian-dollar amounts used in Study 2.

performance. Second, the participants were quite risk-averse, though slightly less so than in Study 1. As indicated in Table 3, the average risk-aversion level was 5.72 with 76.6% exhibiting some degree of risk aversion as detailed in Table 2. Of those remaining, 10% were risk-neutral, while another 13.4% were risk-loving. Third, most participants (50 out of 60, i.e., 83.3%) reported a higher stress level working under the *pfp* setting than the *fs* setting. On a scale of 1 to 5 for level of stress, the mean stress level under *pfp* was 4.48, while it was only 2.28 under *fs*. The mean within-person difference in stress levels between the two compensation schemes of 2.20 was significant ($t = 9.66$, $df = 59$, $p < 0.001$). Fourth, the correlation table (Table 3) shows that the incentive effect of *pfp*, measured by the within-person productivity improvement from the *fs* to the *pfp* rounds, was significantly correlated with individual risk-aversion levels, the individual differences in the variance of performance between the *pfp* and *fs* rounds, the difference in individual reported stress levels between the two pay schemes, and gender. Moreover, while individual risk-aversion levels were positively and significantly correlated with within-person stress differences between *pfp* and *fs*, risk-aversion levels and within-person performance variance differences between the two schemes were negatively and significantly correlated.

Reexamining Hypothesis 1 using the data from Study 2 lends it further strong support. As reported in the first column of Table 4, controlling for gender and country of birth, productivity improvement is once again inversely and significantly related to individual levels of risk aversion ($p = 0.00$). Since we center risk aversion at its mean, the constant term represents productivity improvement for females born in Canada at the sample mean level of risk aversion. The 24 males improved significantly more than the 36 females by 3.72 correct solutions, controlling for risk aversion and place of birth.¹⁵ Similarly, the 17 participants born abroad improved by 2.27 fewer correct solutions than the 43 participants born in Canada, controlling for risk aversion and gender, though this result is of only marginal statistical significance.¹⁶ Since there were no significant interactions between risk aversion and the two control variables, these interactions were dropped from the analysis. Their lack of significance implies that the observed gender effects do not differ between those born in Canada and those born abroad and that the observed country-of-birth effects do not differ between males and females. Most importantly, the negative impact of risk aversion on productivity improvement, the focus of Hypothesis 1, does not differ based either on gender or

¹⁵ Without controls, males improved by 5.71 problems, while females improved by 1.50 problems.

¹⁶ Without controls, those born abroad improved by 1.47 problems, while those born in Canada improved by 3.86 problems.

place of birth.

It is interesting to note that the observed demographic differences arise solely from performance choices made under the *fs* scheme. For example, females actually solve slightly more problems than males under *pfp* (30.22 versus 29.71). However, females do much better than males under *fs* (28.72 versus 24.00). Apparently, on average, males are more inclined than females to choose lower levels of effort under *fs*, resulting in the observed higher levels of improvement under *pfp*. Similarly, those born abroad, mostly in China, solve many more problems under *pfp* than those born in Canada (34.29 versus 28.12). However, this difference widens under *fs* (32.82 versus 24.26). On average, those born abroad seem to exert high levels of effort solving addition problems even in the absence of financial incentives, whereas those born in Canada need financial incentives to perform at high levels of effort. It should be stressed that these demographic observations are based on demographic averages and should not obscure the relationship between risk aversion and performance improvement among all of the demographic groups in our sample.

Hypothesis 2 is predicated on the condition that financial risk is positively correlated with effort as proxied by the number of attempts, N . Let q = the probability of success for each attempt. Note from Table 3 that the actual correlation between q and N is -0.07 , which is not significantly different from zero. Under the corresponding assumption that q is independent of N , output follows a Binomial Distribution with mean = $q \cdot N$ and variance = $q(1 - q) \cdot N$.¹⁷ Since $q(1 - q) > 0$, the variance of output, proportional to financial risk under *pfp*, increases with the number of attempts, our proxy for effort. We cannot directly observe the *ex ante* variance for each participant. However, since each participant played four rounds under *pfp*, we can observe the *ex-post* performance variance for each participant under that payment scheme. This *ex-post* measure can be employed as a proxy for the *ex-ante* variance. A regression of this *ex-post* performance variance for each participant on that participant's number of attempts under *pfp* is reported in column 2 of Table 4.¹⁸ The results confirm a significant positive relationship ($p = 0.05$). Thus, the required condition for H2 to hold is satisfied.

Part (a) of H2 predicts that there should be a significant inverse relationship between a

¹⁷ A discrete binomial distribution may be approximated by the continuous normal distribution employed to derive the theoretical predictions. The normal approximation possesses the same mean and variance as the binomial distribution it approximates with the approximation becoming more accurate as the number of attempts increases.

¹⁸ Initially, dummy variables controlling for gender and place of birth were included in this equation. Since they both were insignificant with p -values greater than 0.30, they were subsequently dropped. In the estimations that follow, both place-of-birth and gender dummies were initially included, but subsequently dropped if they produced p -values greater than 0.30.

participant's degree of risk aversion and both his/her effort and his/her risk exposure under *pfp*. The results of a regression of effort on risk-aversion under *pfp*, controlling for place of birth, are presented in column 3 of Table 4. They corroborate the predicted inverse relationship between risk aversion and effort under *pfp* ($p = 0.00$). Again using the *ex-post* performance variance as a proxy for financial risk exposure, the results reported in column 4 of Table 4 indicate that the predicted significant inverse relationship ($p = 0.02$) occurs in the *pfp* data. As predicted, a similar relationship does not occur in the *fs* data (see column 5) since performance variance is not indicative of financial risk exposure under the *fs* scheme. Accordingly, the difference between the *ex-post* performance variance under *pfp* and the *ex-post* performance variance under *fs* is also significantly and inversely related to the elicited levels of risk aversion as indicated in column 6. Together, these results suggest that participants who are more risk-averse are making a rational decision to undertake less risk exposure under *pfp*, which can be accomplished by investing less effort when money is at stake.

Part (b) of H2 predicts a positive correlation between the productivity improvement from *fs* to *pfp* and risk exposure. This is examined in the first column of Table 5 using a regression of productivity improvement between *pfp* and *fs* on the difference between the *ex-post* performance variance under *pfp* and under *fs*, controlling for gender. Its coefficient is positive and significant as predicted ($p = 0.009$).

Hypothesis 3 is about choking under the stress caused by dislike of payoff uncertainty. The last column of Table 5 reports a regression of the difference in reported levels of stress between *pfp* and *fs* for each participant on individual levels of risk aversion. The increase in a participant's perceived level of stress under *pfp* is positively and significantly related to his/her elicited degree of risk aversion as predicted by H3a. Will the greater increase in stress perceived by more risk-averse participants adversely affect their performance under *pfp* as predicted in H3b? The regression in the first column of Table 6 gives a positive response to this question ($p = 0.02$), controlling for risk aversion, gender and country of birth. Risk aversion itself continues to be significant, indicating that stress is only part of the story. Thus, in addition to the factors discussed in H2, stress partially explains the adverse effect of risk aversion on productivity improvement under *pfp*. More risk-averse participants perceive a greater increase in stress than those who are less risk-averse, and this appears to hinder their improvement in performance.

The regression reported in the first column of Table 6 also indicates that there is a significant interaction between risk aversion and stress ($p = 0.02$) with a negative coefficient. This

implies that the adverse impact of high levels of stress on productivity improvement under *pfp* is stronger, the higher is the level of risk aversion. In order to examine this apparent interaction further, we divided our sample into two subsamples: the 39 participants who improved under *pfp* and the 21 who did not. The results, reported in Table 6, indicate that while risk-aversion was negative and significant for both subgroups ($p = 0.00$ for those who improved and $p = 0.01$ for those who did not), the reported stress difference between *pfp* and *fs* was significant only for those who did not improve ($p = 0.00$).¹⁹ Interaction terms were initially included in these regressions. However, they were insignificant and therefore dropped.

A logit regression was also run to examine Hypothesis 4, which predicts that a higher level of risk-aversion and a greater difference in stress levels will both be inversely related to the probability of an improvement in performance under *pfp*. Also reported in Table 6, it indicates that both independent variables had a negative and significant impact on this probability ($p = 0.01$ for risk aversion and $p = 0.01$ for stress). As there was no significant interaction between these two variables, the interaction term was dropped. These results corroborate H4 and are consistent with our discussion of equation (9), which demonstrated that *ceteris paribus* $r < \tilde{r}$ is consistent with an improvement in performance under *pfp*, while $r > \tilde{r}$ is consistent with a decline in performance. Empirically this relationship is probabilistic because $\tilde{r} = k \cdot \mu_\theta / f \cdot b \cdot \sigma_\theta^2$, which depends on several parameters that are unobservable and will in general differ among participants. Its empirical corroboration suggests that the combination of intrinsic motivation and sufficient risk aversion can result in a perverse response to financial incentives that empirically resembles the crowding out of intrinsic by extrinsic motivation.

In summary, for those who improve under *pfp*, risk aversion adversely affects performance, but not via its impact on stress levels. However, stress levels do adversely affect the probability of being among those who improve. They also significantly affect the deterioration in performance of those who do not improve. Dividing the sample in this manner removes the apparent interaction between risk aversion and stress. It suggests that choking under pressure is a phenomenon affecting

¹⁹ Since this method of dividing the sample was based on the dependent variable, OLS estimates potentially suffer from endogeneity bias. To deal with this possibility, the equations were re-estimated using both a Treatment Effects model and Heckman Selection model. In both cases, the parameters representing the bias (the Inverse Mills Ratio) for each equation were insignificantly different from zero. Thus, OLS is justified in this instance. In any case, the inferences using the Treatment Effects and Heckman Selection models were identical with those under OLS.

only the more risk-averse and stressed-out participants.²⁰

CONCLUSIONS AND DISCUSSION

In this paper, focusing on the role of performance-contingent incentives, we examined the interrelations between individual levels of risk aversion, effort, perceived stress and the effect of financial incentives on performance. We undertook two behavioral studies. The first used an anagram task and Australian participants to demonstrate that improvement in performance under *ppp* is significantly and inversely related to individual levels of risk-aversion. The second demonstrated the robustness of this result using an addition task and Canadian participants. To our knowledge, these studies represent the first laboratory experiments employing salient financial incentives in which attitudes toward risk have been linked with the effectiveness of performance pay at increasing productivity.

Our second study also developed some insights into the mechanisms through which risk aversion affects the improvement in performance under *ppp*. It presents evidence suggesting that more risk-averse people rationally retreat from the risk inherent in *ppp* by choosing to exert less effort even though by doing so they may lower their expected performance and hence their expected pay. Furthermore, we show theoretically and empirically that performance may actually worsen under *ppp*, and that the probability of this occurring is higher, *ceteris paribus*, the more risk-averse a person is. While empirically resembling the crowding out of intrinsic by extrinsic motivation, the underlying reason for this phenomenon is not a loss of intrinsic motivation, but rather a reduction in effort by risk-averse people in order to reduce financial risk.

The second study also shows that more risk-averse people experience a greater increase in stress than less risk-averse people when moving from *fs* to *ppp*. Moreover, this greater increase in stress is associated with a lower probability of exhibiting an improvement in performance under *ppp*. For those who don't improve, the greater the increase in stress, the greater is the deterioration in performance under *ppp*, controlling for demographic factors and the adverse effects of risk aversion on effort and hence performance.

These results are important both in theory and in practice. Theoretically, they suggest that

²⁰ Dividing the sample in this manner is of course somewhat arbitrary. It is possible that some participants who improved slightly under *ppp* also found their performance hindered by the increased stress experienced under that pay scheme, but were too few in number relative to others in their subgroup to have an impact on the overall result for those who improved.

response to financial incentives depends not only on context, but also on individual heterogeneity, and in particular on individual attitudes toward risk. This warrants further theoretical and empirical study, focusing on three important issues.

The first is to examine what mechanisms beyond the specific ones proposed and supported by the data in this paper may lead to a similar inverse relationship between risk aversion and productivity improvement under *ppp*. For example, our theoretical model only allows more risk-averse people to choose less risky strategies through a reduction in effort. For this mechanism to work, effort and risk exposure must be positively correlated. However, such a correlation is not a necessary condition for higher levels of risk aversion to lead to less productivity improvement under *ppp* in real world settings. Any production process that allows each agent to select a level of risk exposure and also exhibits a positive correlation between that risk exposure and expected output can produce a similar result. For example, consider a situation where each agent must select from a set of available projects or work strategies. Assume those projects with greater risk exposures also have higher levels of expected net present value, as in Sung (1995). Effort is chosen independently of risk. Nonetheless, since more risk-averse agents can control risk exposure by choosing less risky projects with lower net present value, an inverse relationship between productivity improvement under *ppp* and risk aversion is quite likely.

Another example of how risk aversion could affect productivity improvement under *ppp* comes from a recent experimental paper by Oswald et al. (2008). Employing the same arithmetic task used in our study, it provides evidence that happier employees are more productive under a piece-rate system. If less risk-averse employees are happier than those who are more risk-averse to work under the uncertainty of *ppp*, this may be another channel, related to but not necessarily the same as stress, through which risk-aversion may affect performance improvement under *ppp*.

A second issue worthy of study is to examine whether the theory presented and corroborated in this paper may help elucidate other empirical findings. For example, an important recent study by Ariely et al. (2009) shows that very high monetary rewards can significantly reduce performance. The authors attribute this effect to stress and choking under pressure. Our results suggest that choking indeed played a vital role. However, it is quite possible that Ariely et al.'s (2009) participants were also rationally choosing to exert less effort in order to reduce risk exposure when exposed to greater financial incentives. Indeed, our theoretical results for piece-rate incentives demonstrate that the critical level of risk aversion, \tilde{r} , falls as the financial incentives rise. This

means that stronger financial incentives are consistent with more people choosing to exert less effort in order to reduce risk. If this was indeed part of the reason for Ariely et al.'s (2009) results, it reinforces their practical importance. While people can become used to performing under pressure and choking can dissipate over time, a rational reduction in effort to reduce risk exposure is more likely to endure.

A third issue concerns stress. Much has been written about stress and performance, primarily in the management and psychology literatures. Some have argued that small amounts of stress may have positive effects on performance, while larger amounts may have negative effects, the so-called “inverted-U” theory (Muse et al., 2003). Others have argued that there are different kinds of stress, in particular “challenge stress”, perceived as “having the potential to promote personal gain or growth, trigger positive emotions and an active problem-solving style” (LePine et al., 2005, p. 765) and “hindrance stress”, perceived as “having the potential to harm personal growth or gain, trigger negative emotions and a passive or emotional style of coping” (LePine et al., 2005, p.765). LePine et al. (2005) link these different types of stress with expectancy theory (Vroom, 1964), suggesting in part that while challenge stress is associated with the belief that there is a positive link between effort exerted and the probability of success, hindrance stress is associated with a more pessimistic belief about the relationship between effort and the probability of success. While acknowledging briefly that “responses to stressors vary somewhat as a function of individual differences that influence the way individuals appraise and cope with stressors” (p. 764), LePine et al. (2005) emphasize that different stressors are usually responsible for the different types of workplace stress. Furthermore, they argue that challenge stressors have a positive impact, while hindrance stressors have a negative impact on performance. We find a significant interaction between risk aversion and stress, with stress levels having no effect on the change in performance of less risk-averse people whose performance improved under *pfp*, but hindering performance of the more risk-averse people whose performance did not improve. This suggests that whether *pfp* is perceived as representing challenge stress or hindrance stress may depend on how much one dislikes financial uncertainty, i.e. on individual levels of risk aversion. For a given perceived link between effort and the probability of success, a more risk-averse person may feel pessimistic and hopeless (“the glass is half empty”), while a less risk-averse person may feel optimistic and hopeful (“the glass is half full”). Thus, we conjecture that individual heterogeneity in risk attitudes may affect one’s perception of whether one is experiencing challenge or hindrance stress under *pfp*, and

thus affect performance. This conjecture deserves further study.

Does self-selection into jobs and compensation schemes lead to people selecting the scheme in which they would be most productive? Dohmen (2008), in a fascinating study of professional football (soccer) players in penalty kick situations finds that choking under pressure increases in front of home crowds, but does not increase in situations for which the stakes are higher. As he points out, the players that specialize in penalty kicks are likely to have self-selected into this task because of their ability to deal well with the resulting stress. However, the fact that many employees self-select into professions and compensation schemes does not imply that all or even most employees have selected the compensation schemes that best match their risk preferences. As Cadsby et al. (2007) demonstrate, while risk aversion is significantly and inversely related to the selection of a *ppp* compensation scheme, it is not the only factor that goes into such a choice. For example, a very risk-averse person may choose a *ppp* scheme if s/he expects to earn a lot more under *ppp* than *fs* due to his/her skill at the work task. However, his/her strong dislike of risk may nonetheless cause him/her to perform worse under *ppp* than under *fs*. S/he can make more money for him/herself under *ppp*, and yet produce less output for the company than under *fs*.

Practically, whether or not employees can self-select into compensation schemes, our results suggest that more risk-averse employees are likely to be less responsive, unresponsive, or even respond in the wrong direction when faced with performance-contingent pay. Since the stakes are bound to be higher in the workplace than in the laboratory, and risk aversion is likely to be greater at higher stake levels (Holt and Laury, 2002), this phenomenon could well be even more pronounced in the workplace than in the laboratory. Thus, different kinds of pay schemes may suit different kinds of workers, and risk attitudes may be a critical factor in determining the best employee-compensation fit.

REFERENCES

- Abowd, John M.** 1990. "Does Performance-Based Managerial Compensation Affect Corporate Performance?" *Industrial and Labor Relations Review*, 43(3): 52–74.
- Ariely, Dan, Uri Gneezy, George Loewenstein, and Nina Mazar.** 2009. "Large Stakes and Big Mistakes." *Review of Economic Studies*, 76(2): 451–469.
- Baker, Ronald J., Susan Laury, and Arlington Williams.** 2008. "Comparing Small-Group and Individual Behavior in Lottery-Choice Experiments." *Southern Economic Journal*, 75(2): 367–382.
- Baker, George P., Michael C. Jensen, and Kevin J. Murphy.** 1988. "Compensation and Incentives: Practice vs. Theory." *Journal of Finance*, 43(3): 593–616.
- Baker, George P., and Bjorn Jorgensen.** 2003. "Volatility, Noise, and Incentives."
http://www.people.hbs.edu/gbaker/oes/papers/Baker_Jorgensen.pdf.
- Banker, Rajiv D., Seok-Young Lee, and Gordon Potter.** 1996. "A Field Study of the Impact of a Performance-Based Incentive Plan." *Journal of Accounting and Economics*, 21(2): 195–226.
- Bénabou, Roland, and Jean Tirole.** 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies*, 70(3): 489–520.
- Baumeister, Roy F.** 1984. "Choking under Pressure: Self-Consciousness and Paradoxical Effects of Incentives on Skillful Performance." *Journal of Personality and Social Psychology*, 46(3): 610–620.
- Baumeister, Roy F., and Carolin J. Showers.** 1986. "A Review of Paradoxical Performance Effects: Choking under Pressure in Sports and Mental Tests." *European Journal of Social Psychology*, 16(4): 361–383.
- Bonner, Sarah E., Reid Hastie, Geoffrey B. Sprinkle, and S. Mark Young.** 2000. "A Review of the Effects of Financial Incentives on Performance in Laboratory Tasks: Implications for Management Accounting." *Journal of Management Accounting Research*, 12(1): 19–57.
- Brase, Gary L.** 2009. "How Different Types of Participant Payments Alter Task Performance."

Judgment and Decision Making, 4(5): 419–428.

Bushman, Robert, Raffi Indjejikian, and Mark Penno. 2000. “Private Pre-Decision Information, Performance Measure Congruity and the Value of Delegation.” *Contemporary Accounting Research*, 17(4): 561–587.

Cadsby, C. Bram, Fei Song, and Francis Tapon. 2007. “Sorting and Incentive Effects of Pay for Performance: An Experimental Investigation.” *Academy of Management Journal*, 50(2): 387–405.

Camerer, Colin, and Robin M. Hogarth. 1999. “The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework.” *Journal of Risk and Uncertainty*, 19(1-3): 7–42.

Deci, Edward L. 1971. “Effects of Externally Mediated Rewards on Intrinsic Motivation.” *Journal of Personality and Social Psychology*, 18(1): 105–115.

Deci, Edward L., and Richard M. Ryan. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Dickinson, David. 2009. “The Effects of Beliefs Versus Risk Attitude on Bargaining Outcomes.” *Theory and Decision*, 66(1): 69–101.

Dohmen, Thomas. 2008. “Do Professionals Choke under Pressure?” *Journal of Economic Behavior and Organization*, 65(3-4): 636–653.

Eriksson, Tor, Sabrina Teyssier, and Marie Claire Villeval. 2009. “Self-Selection and the Efficiency of Tournaments.” *Economic Inquiry*, 47(3): 530–548.

Fehr, Ernst, and Lorenz Goette. 2007. “Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment.” *American Economic Review*, 97(1): 298–317

Fessler, Nicholas J. 2003. “Experimental Evidence on the Links among Monetary Incentives, Task Attractiveness, and Task Performance.” *Journal of Management Accounting Research*, 15(1): 161–182.

Frey, Bruno S., and Reto Jegen. 2001. “Motivation Crowding Theory: A Survey of Empirical

- Evidence.” *Journal of Economic Surveys*, 15(5): 589–611.
- Galbiati, Roberto, and Pietro Vertova.** 2008. “Obligations and Cooperative Behavior in Public Good Games.” *Games and Economic Behavior*, 64(1): 146–170.
- Gerhart, Barry, and George T. Milkovich.** 1990. “Organizational Differences in Managerial Compensation and Financial Performance.” *Academy of Management Journal*, 33(4): 663–691.
- Gerhart, Barry, and Sara Rynes.** 2003. *Compensation: Theory, Evidence, and Strategic Implications*. Thousand Oaks, CA: Sage.
- Gibbons, Robert.** 1998. “Incentives in Organizations.” *Journal of Economic Perspectives*, 12(4): 115–132.
- Gneezy, Uri, and Aldo Rustichini.** 2000. “Pay Enough or Don’t Pay at all.” *Quarterly Journal of Economics*, 115(3): 791–810.
- Gomez-Mejia, Luis R., and Theresa M. Welbourne.** 1988. “Compensation Strategy: An Overview and Future Steps.” *Human Resource Planning*, 11(3): 173–190.
- Hart, Oliver** 1989. “An Economist’s Perspective on the Theory of the Firm.” *Columbia Law Review*, 89(7): 1757–1774.
- Heinemann, Frank, Rosemarie Nagel, and Peter Ockenfels.** 2009. “Measuring Strategic Uncertainty in Coordination Games.” *Review of Economic Studies*, 76(1): 181–221.
- Holmstrom, Bengt.** 1979. “Moral Hazard and Observability.” *Bell Journal of Economics*, 10(1): 74–91.
- Holt, Charles and Susan Laury** 2002. “Risk Aversion and Incentive Effects.” *American Economic Review*, 92(5): 1644–1655.
- Jenkins, Douglas. G. Jr., Atul Mitra, Nina Gupta, and Jason D. Shaw.** 1998. “Are Financial Incentives Related to Performance? A Meta-Analytic Review of Empirical Research.” *Journal of Applied Psychology*, 83(5): 777–787.
- Jensen, Michael C., and William H. Meckling.** 1976. “Theory of the Firm: Managerial Behavior,

- Agency Costs and Ownership Structure.” *Journal of Financial Economics*, 3(4): 305–360.
- Jensen, Michael C., and Kevin J. Murphy.** 1990. “Performance Pay and Top Management Incentives.” *Journal of Political Economy*, 98(2): 225–264.
- Kahn, Lawrence M., and Peter D. Sherer.** 1990. “Contingent Pay and Managerial Performance.” *Industrial and Labor Relations Review*, 43(3): 107–20.
- Kale, Jayant R., Ebru Reis, and Anand Venkateswaran.** 2009. “Rank Order Tournaments and Incentive Alignment: The Effect on Firm Performance.” *Journal of Finance*, 64(3): 1479–1512.
- Kreps, David.** 1997. “Intrinsic Motivation and Extrinsic Incentives.” *American Economic Review*, 87(2): 359–64.
- Lazear, Edward P.** 1986. “Salaries and Piece Rates.” *Journal of Business*, 59: 405–432.
- Lazear, Edward P.** 2000a. “The Power of Incentives.” *American Economic Review*, 90(2): 410–415.
- Lazear, Edward P.** 2000b. “Performance Pay and Productivity.” *American Economic Review*, 90(5): 1346–1361.
- LePine, Jeffrey A., Nathan P. Podsakoff, and Marcie A. LePine.** 2005. “A Meta-Analytic Test of the Challenge Stressor–Hindrancer Stressor Framework: An Explanation for Inconsistent Relationships Among Stressors and Performance.” *Academy of Management Journal*, 48(5): 764–775.
- Locke, Edwin A., and Gary P. Latham.** 1990. *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice Hall.
- Milgrom, Paul, and John Roberts.** 1992. *Economics, Organization and Management*. Englewood Cliffs, NJ: Prentice Hall.
- Muse, Lori A., Stanley G. Harris, and Hubert S. Feild.** 2003. “Has the Inverted-U Theory of Stress and Job Performance Had a Fair Test?” *Human Performance*, 16(4): 349–364.
- Niederle, Muriel, and Lise Vesterlund.** 2007. “Do Women Shy away from Competition? Do Men

- Compete too much?" *Quarterly Journal of Economics*, 122(3): 1067–1101.
- Oswald, Andrew J., Eugenio Proto, and Daniel Sgroi.** 2008. "Happiness and Productivity." Working Paper: Warwick Economic Research Papers No. 882.
- Paarsch, Harry J., and Bruce S. Shearer.** 2009. "The Response to Incentives and Contractual Efficiency: Evidence from a Field Experiment." *European Economic Review*, 53(5): 481-494.
- Paarsch, Harry J., and Bruce S. Shearer.** 2007. "Do Women React Differently to Incentives? Evidence from Experimental Data and Payroll Records." *European Economic Review*, 51(7): 1682–1707.
- Paarsch, Harry J., and Bruce S. Shearer.** 2000. "Piece Rates, Fixed Wages, and Incentive Effects: Statistical Evidence from Payroll Records." *International Economic Review*, 41(1): 59–92.
- Paarsch, Harry J., and Bruce S. Shearer.** 1999. "The Response of Worker Effort to Piece Rates: Evidence from the British Columbia Tree-Planting Industry." *Journal of Human Resources*, 34(4): 643–667.
- Parker, Donald F., and Thomas A. DeCotiis.** 1983. "Organizational Determinants of Job Stress." *Organizational Behavior and Human Performance*, 32(2): 160–177.
- Prendergast, Canice.** 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature*, 37(1): 7–63.
- Rynes, Sara, Barry Gerhart, and Laura Parks.** 2005. "Personnel Psychology: Performance Evaluation and Pay for Performance." *Annual Review of Psychology*, 56: 571–600.
- Seiler, Eric.** 1984. "Piece Rate versus Time Rate: The Effects of Incentives on Earnings." *Review of Economics and Statistics*, 66(3): 363–376.
- Shearer, Bruce S.** 2004. "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *Review of Economic Studies*, 71(2): 513–534.
- Schuler, Randall S.** 1980. "Definition and Conceptualization of Stress in Organizations." *Organizational Behavior and Human Decision Processes*, 25(2): 184–215.

- Schweitzer, Maurice, Lisa Ordóñez, and Bambi Douma.** 2004. "Goal-Setting as a Motivator of Unethical Behavior." *Academy of Management Journal*, 47(3): 422–433.
- Sloof, Randolph and C. Mirjam van Praag.** 2008. "Performance Measurement, Expectancy and Agency Theory: An Experimental Study." *Journal of Economic Behavior and Organization*, 67(3-4): 794–809.
- Stajkovic, Alexander D., and Fred Luthans.** 2001. "Differential Effects of Incentive Motivators on Work Performance." *Academy of Management Journal*, 44(3): 580–590.
- Sung, Jaeyoung.** 1995. "Linearity with Project Selection and Controllable Diffusion Rate in Continuous-Time Principal-Agent Problems." *RAND Journal of Economics*, 26(4): 720–743.
- Vance, Robert J., and Adrienne Colella.** 1990. "Effects of Two Types of Feedback on Goal Acceptance and Personal Goals." *Journal of Applied Psychology*, 75(1): 68–77.
- Vroom, Victor H.** 1964. *Work and Motivation*. New York: Wiley.

FIGURE 1

Hypotheses 1, 2, and 3

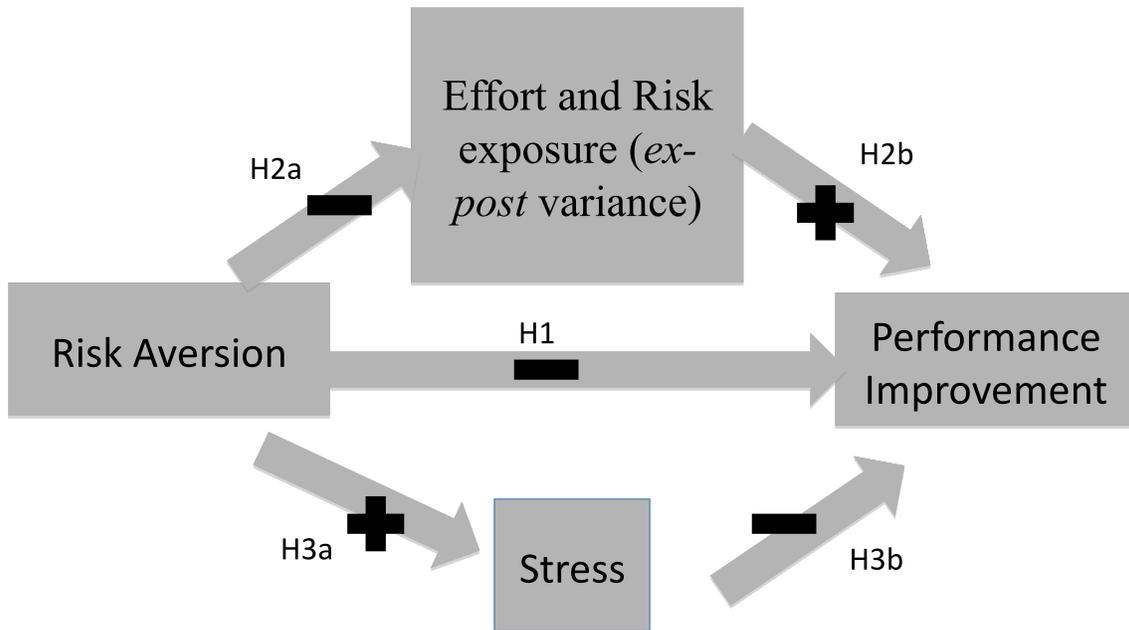


TABLE 1**Study 1: Means, Standard Deviations, and Correlations^a**

	Mean	SD	1	2	3	4	5
1. <i>pfp</i> Output	21.12	5.78					
2. <i>fs</i> Output	18.86	5.70	0.78***				
3. <i>pfp-fs</i> Output Improvement	2.26	3.86	0.35***	-0.2**			
4. Risk Aversion	6.77	1.93	-0.11	0.04	-0.22**		
5. Male Dummy (male=1)	0.62	0.49	0.15	0.21*	0.09	-0.06	
6. Native Language Dummy (English=1)	0.26	0.44	-0.30**	-0.21*	-0.14	-0.07	-0.10

^aN = 115. ***significant at $p < .001$; **significant at $p < .01$; and *significant at $p < .05$ (two-tail test).

TABLE 2**Study 1: Risk-Aversion Classifications Based On Holt-Laury Lottery Choices**

Number of Safe Choices	Risk Preference Classifications (Adopted from HL)	Proportion of Choices	
		Study 1	Study 2
0-1	highly risk loving	3.5%	1.7%
2	very risk loving	0%	5.0%
3	risk loving	0%	6.7%
4	risk neutral	3.5%	10.0%
5	slightly risk averse	9.6%	18.3%
6	risk averse	25.2%	13.3%
7	very risk averse	22.6%	35.0%
8	highly risk averse	21.7%	6.7%
9-10	stay in bed	13.9%	3.3%

TABLE 3

Study 2: Means, Standard Deviations, and Correlations^a

	Mean	SD	1	2	3	4	5	6	7	8	9
1. <i>pfp</i> Output	30.02	9.39									
2. <i>fs</i> Output	26.84	8.97	0.72***								
3. <i>pfp-fs</i> Output Improvement	3.18	6.71	0.42**	-0.32*							
4. <i>pfp-fs</i> Output Variance Diff.	0.02	1.74	0.24	0.03	0.30*						
5. <i>pfp</i> Effort	35.07	9.94	0.94***	0.74***	0.31*	0.26**					
6. <i>pfp</i> Success Rate	0.86	0.09	0.38**	0.13	0.36**	0.003	-0.07				
7. Risk Aversion	5.72	1.81	-0.60***	-0.14	-0.66***	-0.40**	-0.59***	-0.13			
8. Stress	2.20	1.64	-0.30*	0.01	-0.40***	-0.24	-0.29*	-0.14	0.31*		
9. Male Dummy (Male=1)	0.60	0.49	0.03	-0.20	0.31*	-0.06	-0.05	0.23	-0.02	0.02	
10. Born-Abroad Dummy (Born-Abroad =1)	0.28	0.46	0.29*	0.43**	-0.18	-0.05	-0.24	0.20	-0.05	-0.19	-0.14

^aN = 60. *** are significant at $p < .001$; ** are significant at $p < .01$; and * are significant at $p < .05$ (two-tail test).

TABLE 4

**Study 2: The Effects of Risk Aversion on Productivity Improvement and Variance
(*p*-values in parentheses)**

	<i>pfp-fs</i> Output Improvement	<i>pfp</i> Performance Variance	<i>pfp</i> Effort	<i>Pfp</i> Performance Variance	<i>fs</i> Performance Variance	<i>pfp-fs</i> Variance Diff.
Constant	2.36 (0.011)	0.441 (0.502)	33.76 (0.000)	1.725 (0.000)	1.667 (0.000)	0.027 (0.899)
Male Dummy	3.72 (0.004)	Dropped	Dropped	-0.403 (0.262)	Dropped	Dropped
Born-Abroad Dummy	-2.27 (0.098)	Dropped	4.68 (0.04)	0.458 (0.241)	Dropped	Dropped
Centred Risk Aversion	-2.43 (0.000)		-3.16 (0.000)	-0.229 (0.020)	0.155 (0.183)	-0.387 (0.001)
<i>pfp</i> Effort (number of attempts)		0.036 (0.051)				
Adjusted R ²	0.516	0.048	0.369	0.092	0.014	0.148

TABLE 5

**Study 2: The Relationship between *pfp-fs* Output Improvement and *pfp-fs* Variance
Difference and between *pfp-fs* Stress Difference and Risk Aversion (*p*-values in parentheses)**

	<i>pfp-fs</i> Output Improvement	<i>pfp-fs</i> Stress Diff.
Constant	2.023 (0.091)	2.380 (0.000)
Male Dummy	4.240 (0.012)	Dropped
Born-Abroad Dummy	-1.996 (0.263)	-0.641 (0.159)
<i>pfp-fs</i> Variance Difference	1.241 (0.009)	
Centred Risk Aversion		0.273 (0.018)
Adjusted R ²	0.172	0.096

TABLE 6**Study 2: The Effects of Risk Aversion and Stress on *pfp-fs* Output Improvement
(*p*-values in parentheses)**

	All Data (n=60)	Those who improved (n=39)	Those who didn't improve (n=21)	Logit Regression (n=60)
Constant	2.941 (0.001)	5.343 (0.000)	-1.048 (0.1328)	1.026 (0.072)
Male Dummy	3.765 (0.000)	2.260 (0.074)	1.585 (0.253)	1.697 (0.043)
Born-Abroad Dummy	-3.103 (0.015)	-1.702 (0.246)	-1.632 (0.171)	-1.764 (0.038)
Centred Stress Diff	-0.884 (0.019)	-0.220 (0.657)	-1.124 (0.001)	-0.810 (0.013)
Centred Risk Aversion	-2.293 (0.000)	-1.724 (0.000)	-1.181 (0.009)	-0.807 (0.007)
Interaction	-0.407 (0.019)	Dropped	Dropped	Dropped
Adjusted R ²	0.610	0.382	0.572	

Note: Both Risk Aversion and Stress are centered at the mean of the entire sample of 60 participants and the interaction term is the product of these two terms.

Appendix

Instructions for Study 1

(Only data from Rounds 3, 4, 5, and 6 were used in this paper)

Thank you for participating today.

All of your responses in this study will remain completely anonymous. It is important that during this experiment you do not talk or make any noise that might disrupt others around you. If you have any questions, please raise your hand and the experimenter will answer your questions individually.

During this experiment you will be asked to create words using a list of 7 letters. There will be 1 practice round followed by 8 experimental rounds in which you will be given a list of 7 letters to create words. You will have a *Workbook* that will contain all of your work. To ensure anonymity, just write down your participant number on the cover of the *Workbook*. Please do not write your name on any of these materials. Your earnings in this experiment will depend on your performance and/or the specific compensation method applied to each round.

In creating your words, we will use the following rules:

1. It is an English word.
2. It is two or more letters long.
3. It is not a proper noun (e.g., words cannot be names or places).
4. It is made by using each of the 7 letters only once per word (e.g., if the list of 7 letters contains only one ‘g’, you cannot spell “egg”).
5. It is used in only one form (e.g., you cannot use singular and plural versions of the same word).

Here is an example. Say there are 7 letters: SADFTIB. Some examples of permissible words include: “daft”, “fit”, “fad”, “bit”, and “it”. However, the word “dad” is not permissible because the letter “d” is used twice in that word. You may use either “bit” or “bits”, but not both. “I” is not permissible because it contains fewer than 2 letters. If you are unsure if a word conforms to these rules, write it down anyway.

In each round you will create words using the 7 letters at the top of the page in the *Workbook*. You will have 3 minutes to work on each round.

Once we begin the experiment, you will not be able to look ahead to future pages or to go back to previous pages. After the last round, please fill out a short questionnaire. Please respond to the questionnaire as truthfully and as accurately as possible. The questionnaire provides the experimenter with important data. Your responses to the questionnaire are confidential and will not be revealed to anyone other than the experimenter. The data will only be identified by the participant code assigned to you and will not at any point be connected to your name or face in any way.

After you have completed the questionnaire, please raise your hand and the experimenter will escort you to another room where you will be paid your earnings.

Please make sure that you completely understand the instructions for the experiment. Once again, remember not to make any noises that might disturb others around you. If you have any questions, raise your hand and we will answer your questions individually.

Instructions for Study 2

Thank you for participating today. All of your responses in this study will remain completely anonymous. It is important that during this experiment you do not talk or make any noise that might disrupt others around you. If you have any questions, please raise your hand and the experimenter will answer your questions individually.

During this experiment you will be asked to add up sets of five double-digit integers. There will be nine rounds in which you will be given a number of such sets of five integers. The first round is a trial round for you to get familiar with the task while the remaining eight rounds are experimental rounds, which will be used to calculate your earnings. You are not allowed to use a calculator, but you may write numbers down on scratch paper provided by us. The numbers are randomly drawn and each problem is presented in the following way.

98	42	69	50	78	
----	----	----	----	----	--

You will have a *Workbook* that will contain all of your work. Your job is to solve as many problems as you can in each round. Each round lasts 3 minutes. Your earnings in this experiment will depend on your performance and/or the specific compensation method applied to each of the eight experimental rounds. Once we begin the experiment, you will not be able to look ahead to future pages or to go back to previous pages. To ensure anonymity, please write down only your participant number on each page of the *Workbook*. Please do not write your name on any of these materials.

After the last round, please fill out the short questionnaire provided. Please respond to the questionnaire as truthfully and as accurately as possible. The questionnaire provides the experimenter with important data. Your responses to the questionnaire are confidential and will not be revealed to anyone other than the experimenter. The data will only be identified by the participant code assigned to you and will not at any point be connected to your name or face in any way. After you have completed the questionnaire, please raise your hand and the experimenter will escort you to another room where you will be privately paid your earnings.

Please make sure that you completely understand the instructions for the experiment. Once again, remember not to make any noises that might disturb others around you. If you have any questions, raise your hand and we will answer your questions individually.