

Performance Pay and Low-Grade Stress – An Experimental Study

Abstract

BACKGROUND: Although recent economics literature suggests a link between performance pay (PRP) and ill health, this finding is contested on the grounds that this link is plagued by endogeneity between the two variables of interest. **OBJECTIVE:** This study investigates the adverse effects of performance pay on stress which is an important determinant of physical health. **METHODS:** Forty subjects were randomly assigned to two equal groups; either being paid by performance or being paid a flat fee. Both objective (saliva samples to measure cortisol elevation) and subjective (self-reported stress level) measures of stress were obtained before and after participation in the experiment. This experimental methodology purges the effects of self-selection into performance pay and identifies the direction of causation from performance pay to stress which is measured by cortisol levels. **RESULTS:** Those who were paid for their performance experienced higher levels of stress, both in terms of perceived stress and in terms of objectively measured cortisol levels, compared to those who were paid a flat fee for minimum performance. **CONCLUSIONS:** Performance related pay induces objectively measurable stress. Self-reported stress levels and the objective stress measure obtained by measuring cortisol move in a similar direction for the PRP and nonPRP groups, but only the cortisol group shows statistically significant differences between the PRP and nonPRP. This also suggests that individuals are underestimating the stress due to performance pay.

Keywords: Performance-related pay, real-effort experiment, stress, cortisol

Performance Pay and Low Grade Stress – An Experimental Study

1. Introduction

It is long advocated by economists that pay according to performance is the most efficient of the payment schemes [1]. There is a varying incidence of performance-related pay (PRP) in the labour markets of developed countries, which depends on the industry or occupation and on how one defines performance pay. Estimates suggest that it is more than 10-15% of European workers and can be as high as 40% in Scandinavia and the United States (US) [2]. Mainly the literature on performance pay has focused on the effects of such payment schemes on worker productivity. Yet, there are some studies that investigate the interrelationships between working conditions and wellbeing measures such as job satisfaction. Job satisfaction is found to be higher among higher paid workers paid by performance [3, 4] while performance pay is correlated with lower job satisfaction for lower paid workers [4]. One would expect that if performance pay has effects on workers' well-being then it should also have repercussions on the worker's physical health. Indeed, as early as 1776, Adam Smith observed in the *Wealth of Nations*, Book VII that "Workmen ... when they are liberally paid by the piece, are very apt to overwork themselves, and to ruin their health and constitution in a few years".

The focus of this paper is to examine this link between PRP and health (measured by stress) in an experimental setting. As discussed in the following section, the paper offers two advances on the literature. First, the experimental setting controls for the potential for endogeneity which identifies the direction of causality between PRP and health. Second, it uses an objective measure of stress, the hormone cortisol, to allow for a more accurate physiological response to PRP. Below the study discusses the potential pathways through which PRP may affect low

grade stress and hence health, the problem of endogeneity in previous studies and the use of cortisol in measuring stress.

2. Background

There are several multidisciplinary strands of literature that give the context for this study. The discussion below examines the pathways through which PRP could impact health low grade stress and health, how previous studies using survey data can be affected by endogeneity bias making it hard to identify the direction of causation and how an objective measure of stress can be utilised to investigate the link.

2.1 Pathways

There is a long research literature in economics and human resource management that examines the effects of PRP, though the focus is typically on the effects of PRP on worker productivity and resulting pay (e.g. [2] for a general discussion and [5], [6] and [7] for examples in different occupations often showing an increase in pay due to PRP). The key aspect of the standard economic theory of PRP is that it gives stronger incentives for workers to work hard since pay is directly tied to performance.

It is this direct link between the payment method, effort and pay that generates several potential pathways through which performance pay may affect the worker's health. First, there may be an incentive to 'work harder, not smarter' and take more risks at work. Hence, there will be an increase of injuries at work as (particularly manual) workers attempt to increase productivity [8, 9, 10, 11].

A second pathway suggests a direct relationship between performance pay and physical health. Economic theory suggests that PRP explicitly changes the trade-off between work and leisure, giving a relatively larger return to time spent in work. Thus, workers are induced by PRP to shift hours to work and away from nonwork activities [12]. Since nonwork activities include time spent on healthy behaviours (such as exercising, sleep, leisure or shopping and cooking healthy meals), an increase of time spent at work should be expected to reduce the above activities. The effects of this change would not manifest themselves on immediate health deterioration (as opposed to having an injury), but their adverse repercussion on health would build up over time. In line with this, the longer the time spent in jobs with performance pay, the higher the odds of having worse overall health, heart problems, stomach problems, and anxiety/depression [13].

A third pathway links performance related pay directly to increases in the individual's exposure to stress and his or her health. For instance, data from the General Health Questionnaire (GHQ) show a strong correlation between time spent in PRP jobs and reported stress level. In addition, the literature suggests that higher job stress is associated with lower job satisfaction and higher injury-related absenteeism [5, 14, 15].

The medical literature (e.g. [16, 17]), provides a well understood link between the exposure to stress and the state of the individual's physical and psychological health. In summary, the physiological processes involved are as follows: In response to stress the immune system redirects white blood cells to areas where injury or infection is most likely, the skin becomes cool and sweaty as blood is drawn away from it toward the heart and muscles which are essential for survival, the mouth becomes dry, and the digestive system slows down. However, when the cause of stress passes, the levels of stress hormones drop and the body's various organ

systems return to normal, a state called 'Allostasis'. However, an absent or incomplete relaxation response may cause damage as 'Chronic Stress' or 'Low Grade Stress' prevents physiological arousal from returning fully to normal. A review of the literature on the effects of acute and chronic stress [18] suggests that although the body's natural defences adapt and thus people can overcome episodes of stress, excessive chronic stress which is constant and persists over an extended period of time can be psychologically and physically damaging. Hence, there is a relationship between inflammatory responses to chronic psychosocial stress and long-term development of disease. The physical effects of chronic stress or 'low grade stress' burden the cardiovascular system, the immune system (increases a person's risk of getting an infectious illness), the brain (interferes with memory and learning), the musculoskeletal system (intensifies the chronic pain of arthritis and produces tension-type headaches) and the reproductive system (can cause impotence in men and affects fertility) [18].

2.2 Endogeneity

In view of these repercussions of chronic stress on health, the focus of this study is to investigate if performance related pay systems significantly increase low grade stress. If this is so then since most of the life of a working person is spent at work, long exposure to PRP should be expected to generate chronic or low-grade stress which can potentially induce severe deterioration of the working person's health. However, the link between PRP systems and health is not straightforward. It is not clear if the causation runs from the payment method to increases in stress (and subsequent deterioration in health) or if poor health tends to drive workers into certain kinds of payment methods such as PRP. If a correlation is found between working on a PRP contract and low health status, this may be an outcome of the propensity of individuals with low health status who being unable to perform and hold a job in regular contracts are dislocated to inferior PRP contracts. Alternatively, those who choose PRP

contracts might be more adept in stressful situations so subjectively care less about their exposure to stress. Thus, one of the biggest problems in identifying econometrically any link between work contracts (payment methods) and health is that there may be endogeneity between two variables of interest [19]. To disentangle these effects, econometric research either relies on Heckman-type corrections [20] for endogenous selection or instrumental variables procedures. In both cases, the unbiasedness of the estimates depends crucially on the statistical properties of the identifying restrictions. Typically, the case for any selection of identifying restrictions to control statistically for this endogeneity can be challenged on statistical or theoretical grounds. The issue could be resolved if workers were randomly distributed in the different payment contracts, but this is typically not an option in real labour markets. Thus, the present study uses experimental methodology to investigate the possible link between PRP and stress by circumventing endogeneity through initial randomisation of the subjects.

2.3 PRP and Low Grade Stress

There are only two papers to date that use an experimental design to look at PRP and stress. The first reports an experiment examining the sorting of more productive players into situations where there are performance-related payments as a result of the game [21]. However, the linkage between stress and PRP is not central to their study, the experimental subjects are simply asked about their stress and exhaustion at the end of the experiment, and those who are in PRP express higher levels of stress and exhaustion at the completion of the experiment.

The second study is an experimental study where they examine whether there is increased productivity due to PRP [22]. They also ask the experimental subjects about levels of stress and find that the increase in productivity induced by PRP is fully offset by the increased stress

of such a payment scheme such that 25 percent of subjects had lower productivity when paid for their performance particularly among the risk averse.

Both the above papers suggest that there is a link between PRP and stress, but both allow selection into PRP and so may be biased by endogeneity. In addition, both use self-reported Likert scale responses to a question inquiring about the stress felt by the subjects during the PRP experience. While self-reported stress may be suggestive of the true underlying, physiological changes in stress, it is not an objective measure of stress as subjective responses are contaminated by beliefs, cognitive dissonance and adaptation effects.

In psychology and more broadly in the medical literature, stress can be measured objectively through assessing the individual's heart rate, blood pressure, or cortisol levels in blood or saliva. Cortisol is a hormone secreted when people are in stressful situations. In surveys of the literature on the use of cortisol as a physiological indicator of stress by [23] and [24], it is found that cortisol is quickly released into the body when a person experiences stress. In addition, the cortisol appears in saliva, thus negating the need for invasive blood tests (which may themselves potentially cause stress in subjects). Furthermore, the saliva test remains accurate even at room temperature for at least a week and hence facilitates repeated experiments.

In addition to being an important signal of stress to the body, the release of cortisol has real effects on the body. Under normal circumstances, cortisol (as part of the hypothalamic-pituitary-adrenal or HPA axis) helps to regulate the body's response to stress by generally suppressing reactions to stress through allostasis, helping the body return to a normal equilibrium [25]. However, repeated or chronic stress can cause 'allostatic load' [26] which

dampens the ability of the body to return to ‘normal’ either by causing stress reactions such as increased blood pressure to continue beyond the direct impact of stress or by suppressing normal (e.g. immune system) responses to stress [26]. While these are simple examples and medical research [27] finds that the mechanisms can be quite complex, current medical research suggests a strong link between cortisol (and other HPA-axis hormones) and adverse health outcomes.

2.4 Summary

The literature above suggests a link between PRP and stress or health. However, problems with endogeneity and subjectivity may be influencing these links. Thus, the present study, first, randomly allocates subjects to two groups to circumvent endogeneity arising from selection bias. In one group, subjects are paid a flat rate while subjects in the other group are paid for their individual performance. Second, it uses an objective marker of stress (salivary cortisol level). The next section details how the design of the experiment accomplishes these two innovations in this literature.

3. The Design of the Experiment

The performance task used in this study was to have subjects calculate a variety of mathematics problems by hand and enter the result in a computer. This is similar to the methodology utilised by [21]. These calculations last for ten minutes which is sufficient time for the rise of cortisol levels in the presence of possible stress. Subjects are randomly assigned either being paid by the number of questions answered correctly (the ‘PRP’ group) or being paid a flat fee for answering ten questions correctly (the ‘nonPRP’ group). During the experiment the computer program ‘z-tree’ [28] is used in order to record the correct answers and calculate payoffs. The study protocol has been reviewed and approved by the University [Edited for Review Process]

Forty subjects (average 22.7 years of age, 57.5% are female) were invited to the Scottish Experimental Economics Lab (SEEL) at the University of [Edited for Review Process] over two sessions (20 subjects per session) to generate the data. The students were recruited using the Online Recruitment System for Economic Experiments (ORSEE) database of potential subjects which is maintained by SEEL. Students were given details about the broad parameters of the experiment and the procedure of the cortisol sample testing. The subjects were also advised, in line with standard cortisol testing protocols, to abstain from eating, drinking caffeine, smoking or taking exercise two hours before the commencement of the experiment. To this effect reminders were sent via email 24 hours before the experiment was scheduled to take place. The two experimental sessions took place at 1500 hours on the Monday and the Wednesday of the same week to control for the known diurnal patterning of cortisol production and to standardise the experience of participants. Saliva samples and subjective stress reports were obtained before and after participation in the experiment. These provide the objective and subjective the stress measures for the individual subject, respectively.

The computer generated randomisation allocated 57.5% (23 out of 40) to the PRP group and the remaining 17 to the flat fee group. The random assignment of PRP to subjects addresses the issues of endogeneity and non-random selection discussed above. Therefore, the experiment tests whether there is a direct causal relationship between PRP and stress level by examining differences in cortisol (and the subjective measures of stress) across the two groups. Thus, one should expect that any comparisons should be representative of the direction and the strength of the PRP – stress relationship directly without any interference of the self-selection effects.

3.1 The Experiment

Upon arrival to SEEL, subjects were given a consent sheet that they signed and were randomly allocated a seat at a computer terminal. Screens between terminals prevented subjects from seeing other subjects and their terminals. When all 20 subjects were registered, an outline of the experiment was read to the group and any questions were answered. All subjects were informed that for their participation they would earn £5, with the opportunity to earn more money during the experiment.

In the next stage a baseline of stress measurement for the subjects was obtained. First, instructions were given for the cortisol test, which involved chewing a cotton swab (SalivaBio Oral Swab, Salimetrics Europe) for 60 seconds and then placing the swab in a test-tube. Subsequently, a computer-based questionnaire was given to each subject, including the item 'how stressed do you feel?' (1=not stressed at all, 5=very stressed) (as in [21]).

After the completion of the questionnaires about stress, all subjects were given a practice round of answering three mathematical questions – one addition (in thousands), one multiplication (hundreds by tens) and one division (thousands by tens). Subjects were allowed to use a scratch piece of paper, but no calculator. This part of activity was not timed. After everyone completed this practice round, the computer program randomly assigned subjects into the PRP or nonPRP group. Subjects were individually told by the computer how they would be paid. The payment schemes were as follows: 20p for each correct answer for those randomly allocated to the PRP group and a £5 flat payment for the nonPRP group if at least ten questions were answered correctly. Subjects worked independently and were not aware of what the alternative payment schedule was or of how other subjects were being paid. Subjects were told that they had ten minutes to do the mathematics questions and that there were a maximum of 50 questions.

When all were ready, a clock appeared on all screens counting down the time in seconds. During the task, the number of questions answered correctly was also displayed at the top of the screen. The mathematical questions (either multiplication, addition, or division) were shown in the middle of the screen. When those in the nonPRP group accomplished the minimum performance target of ten correct answers, this was indicated at the top of the screen. Although they were not eligible to receive any additional pay and were told this on their screen, subjects in this group could still continue to answer questions if they wished.

At the end of ten minutes, the task was stopped. Because the cortisol response to an acute stressor peaks around 20 minutes after stressor onset ([29, 30]), the subjects were asked to leisurely complete several tasks for the 10 minutes after the (10 minute) experiment was complete. First, they again gave a subjective rating of their stress level and provided some demographic information (e.g. gender, year at university, broad discipline of studies and age) before completing two nonstressful filler tasks (rating vignettes of potential jobs with different characteristics and colouring patterns). At the end of the ten minutes, subjects took a second cortisol test by chewing on another swab for one minute and putting the swab in a different, labelled test-tube after the minute was over.

Finally, subjects were called to the control room by seat number to get their payment and were thanked for their participation. Test-tubes were transferred into a freezer and when both sessions were complete, the frozen samples were sent to a laboratory (Salimetrics, Europe) for analysis of cortisol levels for each subject before and after the task.

4. Results

On average, PRP subjects answered 34.7 questions correctly (SD 9.6) whereas the nonPRP answered 32.1 questions correctly (SD 11.4). The standard deviation regarding the correctly answered questions is greater for the nonPRP mainly due to the fact that some subjects stopped after meeting the minimum requirement of correct questions. The average pay-out was £11.90 for the PRP group ranging from £8.80 to £14.60 and £10 for the nonPRP group. It is noteworthy that all nonPRP subjects obtained the minimum required performance.

An initial analysis of the cortisol measures of stress identified one subject whose cortisol measurement in the first assay was assessed to have a hormone level more than four standard deviations from the mean level of cortisol for the other 39 subjects. Given the likelihood that this measure reflected contamination and the potential for this clear outlier to affect comparisons, the information for this particular subject was excluded from the subsequent analysis, leaving 39 subjects in the analysis.

4.1 Stress Results by Group

Table 1 and Figure 1 shows the means of the stress levels for the cortisol and subjective measurements at the start (before PRP assignment) and the end of the experiment. A clear pattern emerges from these results with the objective and subjective measures of stress increasing for the PRP group and falling for the nonPRP group. Figure 2 shows the differences between the start and the end of the experiment for the stress of the two groups for the two measures, again showing that for the PRP group there is some positive elevation in stress reflected in both cortisol and subjective measures of stress but there are noticeable decreases for the two measures in the nonPRP group.

4.2. Differences in Cortisol

While suggestive, these figures do not indicate whether these changes are statistically different. In order to test this, first there is a need to establish that both the PRP and the nonPRP groups have non-significant differences in the levels of mean stress before the treatment. To assess the validity of this requirement a t-test comparing the level of stress for the PRP and nonPRP groups before treatment is employed.

As [31] suggest that unequal variances can influence the analysis. Hence in the first step, the Folded F-test for unequal variances is used to evaluate if equality of variances can be assumed. If there is support for inequality of variances, the Satterthwaite test for unequal variances is utilised, otherwise the Pooled test for equal variances is used. Research [31] suggests that when in doubt it is safer to assume unequal variances.

In view of the above, first the homogeneity of variances between the two groups is assessed using the Folded F-test (see Table 1) that indicate that the two groups may have unequal variances with a p-value of 0.0343. In view of this, the Satterthwaite Unequal Variances test is employed giving a test statistic of 1.90 which shows that the null hypothesis of equality of means cannot be rejected at the 0.05 level of statistical significance.

Therefore, a Two-Sample t-test for the difference of the changes of stress between the start and the end of the experiment for the two groups is employed, under the null that there is no difference in the elevation of stress between the PRP and the non-PRP group. A Folded F-test for homogeneity of variances post-treatment shows that the null cannot be rejected with a p-value of 0.1867. Hence the Pooled for Equal Variances t-test reveals that the differences in the changes of stress between PRP and nonPRP is statistically significant with a p-value of 0.0476.

Given that the stress level of the PRP group went up, then this lends support to the conjecture that the PRP payment scheme increases stress in subjects.

4.3 Differences in Subjective Stress

The above statistical procedure is also used for the subjective stress measure. The Folded F-test for the equal variances before assignment suggest that there is no difference in variances in the subjective stress measure (p-value of 0.3327). Thus, a standard t-test of equal means across the two groups cannot reject the null of equal means of the subjectively measured stress at the start of the experiment (p-value of 0.2068).

For the post-treatment differences in variances, a Folded F-test suggests no difference in variance (p-value of 0.3567). Interestingly, however, the Pooled for Equal Variances test reveals that the differences in the elevation of subjective stress between PRP and nonPRP are statistically insignificant (p-value of 0.2897).

Productivity Measurement: Figure 3 shows that differences in productivity between the PRP and the nonPRP group as reflected in the group mean in correct answers during the experiment. The figure suggests no great difference in productivity between the two groups. This is confirmed by the formal statistical test. The Folded F-test cannot reject (p-value of 0.224) the null of equal variances. Testing the equality of the average number of correct answers, the Pooled for Equal Variances t-test shows that the difference in productivity between the two groups is statistically insignificant (p-value: 0.4515).

5. Discussion

This study provides a real-effort experiment to reveal whether there is a link between PRP and low grade stress. Unlike previous research on stress and PRP in an experimental setting, the experimental methodology in this study used random assignment to PRP to circumvent concerns of self-selection and endogeneity bias and measures stress by both subjective and objective means.

The results show that low grade stress increases over the course of a simulated work task for subjects paid by PRP but falls for the nonPRP group. The differences in stress between the start and the end of the experiment show that for the PRP group there is some positive elevation in stress reflected in both cortisol and subjective measures of stress but there are noticeable decreases for the two measures in the nonPRP group. Furthermore, the study reveals that the differences in the changes of objectively measured stress between PRP and nonPRP are statistically significant. In view that the stress level of the PRP group increased, this lends support to previous studies proposing that the PRP payment scheme is inducing stress in subjects.

However, the study reveals that contrary to the objective measures of stress through cortisol elevation, the small differences in the elevation of subjective stress between PRP and nonPRP are statistically insignificant. This implies that the respondents are not aware of their own body's response to stress. Hence subjective responses may be contaminated by beliefs or cognitive dissonance. This issue needs further investigation. Finally, the results show that contrary to several studies (e.g. [1]), there appear to be no significant difference in productivity between the two groups.

Although the results are suggestive that PRP generates stress, this is a small study based only on 39 respondents. Future work will expand on this sample to measure the effects of PRP on stress across different groups. For example, as [21] and others suggest, there are important differences in the effects of PRP by gender, perhaps caused by attitudes of risk. Although these issues should be circumvented due to the randomisation methodology of this study, further investigation for these issues is important. Another interesting extension moves beyond random assignment which is more similar to a 'real world' change in the employment contract for a firm. Many employment situations involve the element of choice of PRP or nonPRP (e.g. when people are looking for jobs) and the interaction between stress and PRP may be different in these cases. Allowing subjects to choose a PRP or nonPRP job would allow for this selection in an experimental setting and an insight on whether selection is actually a mitigating influence on stress in PRP settings – a common critique of survey-based papers on the stress-PRP relationship.

6. Conclusions

Results suggest that the subjective measure of self-reported stress levels and the objective stress measure obtained by measuring cortisol move in a similar direction for the PRP and nonPRP groups, but only the movement in the cortisol shows statistically significant differences between the two groups. This might suggest that individuals are underestimating the stress generated by the PRP method of payment systems, though more work will be needed to test this formally.

Acknowledgements: [Edited for Review Process]

REFERENCES:

- [1] Lazear EP. Salaries and piece rates. *J Bus.* 1986; 59:405-431.
- [2] Bryson A, Freeman R, Lucifora C, Pellizzari M, Perotin V. Paying for performance: incentive pay schemes and employees' financial participation. In: Boeri T, Lucifora C, Murphy KJ, editors. *Executive remuneration and employee performance-related pay: A transatlantic perspective.* Oxford: Oxford University Press; 2013. p.123.
- [3] Green CP, Heywood JS. Does performance pay increase job satisfaction? *Economica.* 2008; 75(300): 710-728.
- [4] McCausland WD, Pouliakas K, Theodossiou I. Some are punished and some are rewarded: A study of the impact of performance pay on job satisfaction. *Int J Manpower.* 2005; 26(7/8): 636-659.
- [5] Adaramola, SS. Job stress and productivity increase. *Work.* 2012, 41, 2955-2958.
- [6] Lavy V. Performance pay and teachers' effort, productivity, and grading ethics. *Am Econ Rev.* 2009; 99: 1979-2011.
- [7] Liu X, Mills A. The effect of performance-related pay of hospital doctors on hospital behaviour: A case study from Shandong, China. *Hum Resourc Health.* 2005; 3(11).
- [8] Freeman RB, Kleiner M. The last American shoe manufacturers: decreasing productivity and increasing profits in a shift from piece rates to continuous flow production. *Ind Relat.* 2005; 44(2): 307-330.
- [9] Bender KA, Green CP, Heywood JS. Piece rates and workplace injury: does survey evidence support Adam Smith?" *J Popul Econ.* 2012; 25(2): 569-590.
- [10] Artz B, Heywood JS. Performance pay and workplace injury: panel evidence. *Economica.* 2015; 82(S1): 1241-1260.
- [11] DeVaro J, Heywood JS. Performance pay and work-related health problems: A longitudinal study of establishments. *Ind Labor Relat Rev.* 2017; 70(3); 670-703.
- [12] White M, Hill S, McGovern P, Mills C, Smeaton, D. 'High-performance' management practices, working hours and work life balance. *Brit J Ind Relat.* 2003; 41(2); 175-95.
- [13] Bender KA, Theodossiou I. The unintended consequences of the rat race: the detrimental effects of performance pay on health. *Oxford Econ Pap.* 2014; 66(3): 824-847.
- [14] Heo YS, Leem JH, Park SG, Jung DY, Kim HC. Job stress as a risk factor for absences among manual workers: a 12-month follow-up study. *Ind Health.* 2015, 53, 542–552.
- [15] Drakopoulos S, Grimani K. Injury-related absenteeism and job satisfaction: insights from Greek and UK Data. In *J Hum Res Manag.* 2013, 24: 3496– 3511.

- [16] Jansen AS, Nguyen XV, Karpitsky V, Mettenleiter TC, Loewy AD. Central command neurons of the sympathetic nervous system: basis of the fight-or-flight response. *Science*. 1995; 270(5236): 644-646.
- [17] Gleitman H, Fridlund AJ, Reisberg D. *Psychology*. London: WW Norton & Company; 2004.
- [18] Rohleder N. Stimulation of systemic low-grade inflammation by psychosocial stress. *Psychosom Me*. 2014; 76(3): 181-189.
- [19] Eriksson T. Healthy personnel policies. *Int J Manpower*. 2012; 33(3): 233-245.
- [20] Heckman JJ. Sample selection bias as a specification error. *Econometrica*. 1979; 47: 153-161.
- [21] Dohmen T, Falk A. Performance pay and multidimensional sorting: productivity, preferences, and gender. *Am Econ Rev*. 2011; 101(2): 556-590.
- [22] Cadsby CB, Song F, Tapon F. The impact of risk-aversion and stress on the incentive effect of performance pay. In: George SJ, Hamman JR, editors. *Experiments in Organizational Economics (Research in Experimental Economics, Vol 19)*. London: Emerald Group Publishing Limited; 2016; p. 227.
- [23] Kirschbaum C, Hellhammer DH. Salivary cortisol in psychobiological research: an overview. *Neuropsychobiology*. 1989; 22(3): 150-159.
- [24] Nicolson N. Measurement of cortisol. In: Luecken L, Gallo L, editors. *Handbook of physiological research methods in health psychology*. Thousand Oaks, CA: Sage Publications; 2008; p. 37.
- [25] McEwen BS. Stressed or stressed out: what is the difference? *J Psychiatr Neurosci*. 2005; 30(5): 315-318.
- [26] McEwen BS. Protective and damaging effects of stress mediators. *N Engl J Med*. 1998; 338(3): 171-179.
- [27] Miller GE, Chen E, Zhou ES. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychol Bull*. 2007; 133(1): 25-45.
- [28] Fischbacher U. Z-Tree: Zurich toolbox for ready-made economic experiments. *Exp Econ*. 2007; 10(2): 171-178.
- [29] Kirschbaum C, Pirke KM, Hellhammer DH. The 'Trier Social Stress Test' – A tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*. 1993; 28(1-2): 76-81.
- [30] Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull*. 2004; 130(3): 355-391.

[31] Snedecor GW, Cochran WG. Statistical Methods. Ames, Iowa: Blackwell Publishing Professional; 1989.

Table 1. Descriptive results of stress by payment type.

	Self-reported stress (1-5 likert scale)		Cortisol measurement	
	Before Task	After Task	Before Task	After Task
PRP group	2.68	3.00	0.104	0.108
NonPRP group	2.29	2.24	0.159	0.121

Table 2. Statistical Tests Applied

Test	Type of test	Cortisol Measurement		Subjective Stress Measurement	
		Degrees of Freedom	Test statistic (p-value)	Degrees of Freedom	Test statistic (p-value)
Equal variances before assignment	Folded F-test	16, 21	2.35 (0.0343)	16, 21	1.21 (0.3327)
Equality of differences in change in stress	Satterthwaite Unequal Variances test	26.1	1.90 (0.0679)		
Equality of means before assignment	t-test for mean difference			37	-1.28 (0.2068)
Equal variances after treatment	Folded F-test	21,16	1.55 (0.1867)	21,16	1.20 (0.3567)
Equality of differences in change in stress	Pooled for Equal Variances t-test	37	2.11 (0.0476)	37	1.07 (0.2897)

Figure 1. Cortisol and Subjective Stress Measurements; Before and After Treatment

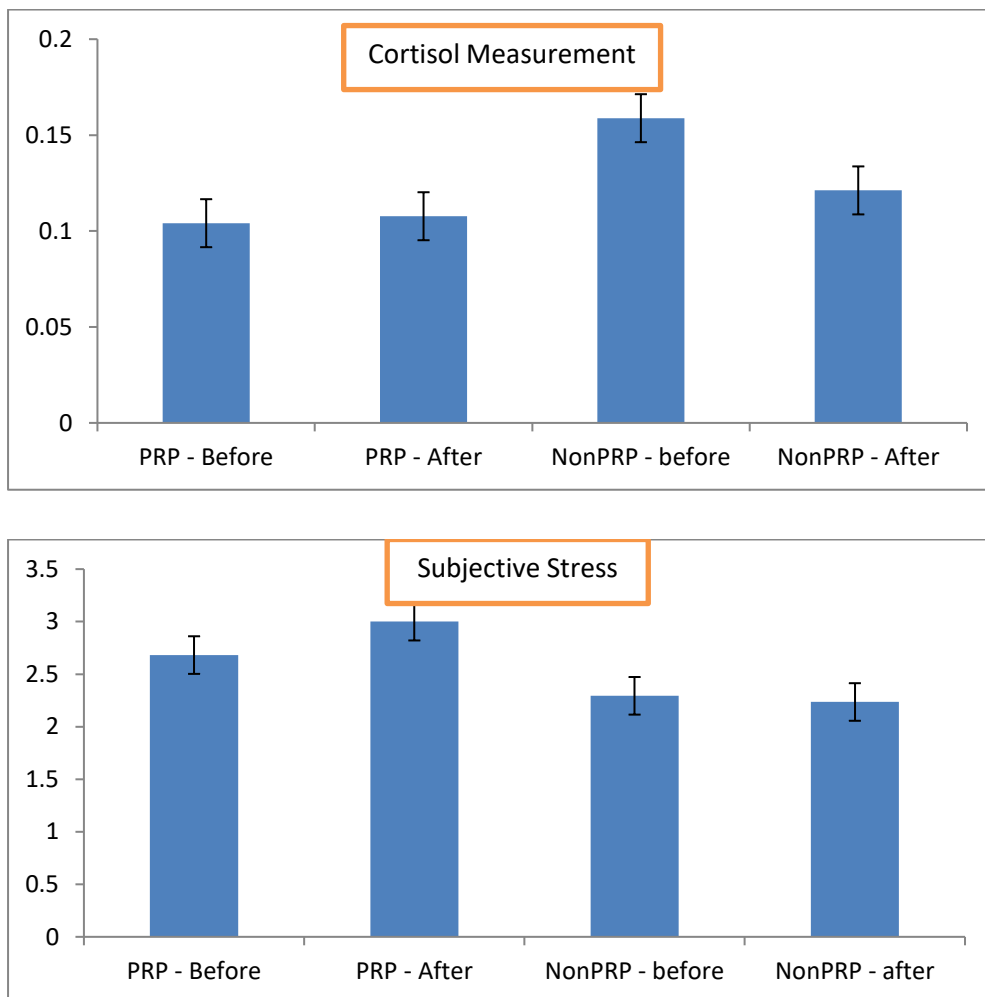


Figure 2: Cortisol and Subjective Measurements; Before and After Treatment

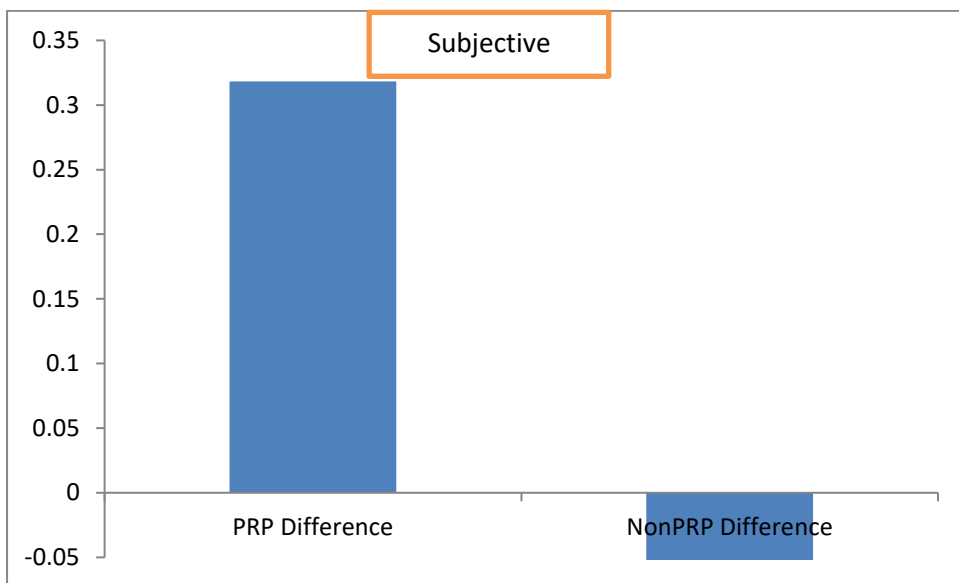
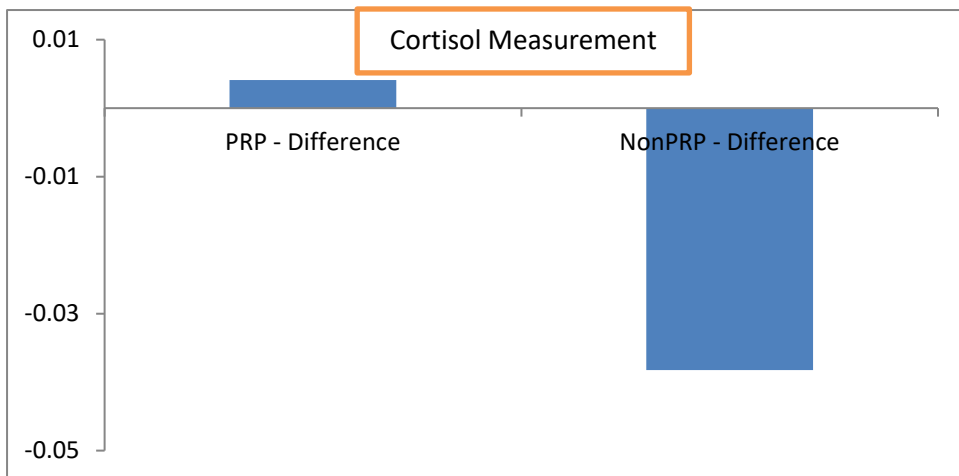


Figure 3. Productivity Measurement; Correct Answers.

