Experimental Economics (2023) 26:955–985
https://doi.org/10.1007/s10683-023-09801-1

ORIGINAL PAPER

Tournament incentives affect perceived stress and hormonal stress responses

Thomas Dohmen[1,2,3] · **Ingrid M. T. Rohde**[4] · **Tom Stolp**[3,5]

Received: 27 January 2022 / Revised: 23 March 2023 / Accepted: 24 March 2023 /
Published online: 26 April 2023
© The Author(s) 2023

Abstract

We conduct a laboratory experiment among male participants to investigate whether rewarding schemes that depend on work performance—in particular, tournament incentives—induce more stress than schemes that are independent of performance—fixed payment scheme. Stress is measured over the entire course of the experiment at both the hormonal and psychological level. Hormonal stress responses are captured by measuring salivary cortisol levels. Psychological stress responses are measured by self-reported feelings of stress and primary appraisals. We find that tournament incentives induce a stress response whereas a fixed payment does not induce stress. This stress response does not differ significantly across situations in which winners and losers of the tournament are publically announced and situations in which this information remains private. Biological and psychological stress measures are positively correlated, i.e. increased levels of cortisol are associated with stronger feelings of stress. Nevertheless, neither perceived psychological stress nor elevated cortisol levels in a previous tournament predict a subsequent choice between tournaments and fixed payment schemes, indicating that stress induced by incentives schemes is not a relevant criterion for sorting decisions in our experiment. Finally, we find that cortisol levels are severely elevated at the beginning of the experiment, suggesting

✉ Tom Stolp
t.stolp@maastrichtuniversity.nl

Thomas Dohmen
t.dohmen@uni-bonn.de

Ingrid M. T. Rohde
ingrid.rohde@ou.nl

[1] Institute for Applied Microeconomics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany

[2] IZA, Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany

[3] Maastricht University, Tongersestraat 53, 6211 LM Maastricht, The Netherlands

[4] Open Universiteit, Valkenburgerweg 177, 6419 AT Heerlen, The Netherlands

[5] Vrije Universiteit Amsterdam, Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

that participants experience stress in anticipation of the experiment per se, potentially due to uncertainties associated with the unknown lab situation. We call this the novelty effect.

Keywords Incentives · Stress · Cortisol · Sorting · Laboratory experiment

JEL Classification D23 · D87 · D91 · M52

1 Introduction

Working life is increasingly associated with stress. More than a quarter of European workers report that psychosocial stress affects their mental well-being (EU-OSHA, 2013; Eurofund and EU-OSHA, 2014). Not least because work-related stress imposes substantial costs on individuals and society, it is important to better understand which working conditions induce stress.[1] Performance pay—i.e. earnings *depend* on the performance of the worker—is one candidate that potentially contributes to work-related stress. In this paper we test in a laboratory experiment whether tournament incentives, one prototype of performance pay, induces acute stress, the antecedent of chronic stress.[2]

The laboratory setting enables us to accurately measure stress responses to different forms of remuneration, which is difficult to capture using naturally occurring data. Moreover, incentives at the workplace are often confounded as workers face a mix of short-term and long-term monetary and non-monetary incentives, including career concerns. In our laboratory setting we can isolate the effect of incentives on stress by exogenously varying the worker's incentive schemes. This enables us to causally identify whether an incentive scheme induces a stress response. Another complication in real world settings is that observed stress might also be induced by stressors other than incentive schemes. Importantly, measures of acute stress induced by a particular stressor are typically not available in field data, and would be difficult to elicit in most work settings. Recording the hormonal stress reaction, for example, requires the sequential collection of biomarkers over a period of about 50 min. In our lab setting, we can repeatedly elicit both biomarkers and self-reports of stress.

[1] Chronic stress, i.e. exposure to stress over longer periods of time, is detrimental to physiological and psychological health, contributing to cardiovascular disease, metabolic syndrome, diabetes mellitus and mental disorders (see e.g. Cohen et al., 2007; McEwen, 2008). In Europe, for example, the cost of mental health disorders is estimated at EUR 240 billion per year (EU-OSHA, 2013). These figures do not take into account additional individual costs of stress due to reduced quality of life and its burden on caretakers and families. In addition, stress arguably affects the quality of economic decision-making (Mani et al., 2013), and hence is likely to affect behavior across all domains of life. Studies in psychology and psychiatry have documented that anxiety disorders result in reduced capacity of working memory and cognitive task performance (see Eysenck, 1992).

[2] Clearly, our results can only be a first step towards assessing whether performance pay can be a pathway to adverse health externalities. We can only establish whether incentive pay induces an acute stress response, which is a prerequisite for chronic stress. However, we cannot assess in our experiment whether repeated exposure to incentive pay provokes recurrent stress responses. Nevertheless, it is important to answer the questions that are addressed in this paper, as an acute stress response is a prerequisite for negative health consequences induced by performance pay.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

In order to investigate the effect of incentives on stress, we expose participants in our experiment consecutively to different types of remuneration for their performance on a 10-min multiplication task. Each participant performs the task under three different payment conditions: (1) a fixed payment scheme, (2) a two-person tournament in which the winner receives a cash prize and the loser gets nothing, and (3) a choice between the fixed payment scheme and the two-person tournament. The order of exposure to the first two payment conditions is randomized. Since tournament incentives may induce acute stress due to greater uncertainty, or the threat of social evaluation that is inherent in many incentive schemes in which performance is revealed to others, we implemented two treatment conditions that varied whether winners are publicly or privately announced. In the public disclosure treatment, participants form a circle near the end of the experiment and turn-by-turn reveal the tournament outcome to each other. In the private disclosure treatment, participants are not required to reveal the tournament's outcome to each other.

To track individuals' stress reactions that are caused by different incentive schemes, we repeatedly measure salivary cortisol concentration and self-reported stress during each payment condition. Cortisol is a steroid hormone that is released during the physiological stress response along the hypothalamic-pituitary-adrenocortical (HPA) axis and is a well-established objective physiological measure of acute stress (Kirschbaum et al., 1993). Since homeostasis is reached about 50 min after the onset of a stressor, we track cortisol in each payment condition by taking cortisol swaps at the baseline and after 20, 30 and 50 min after the onset of the stressor. Moreover, we take additional saliva samples on two weekdays during the week preceding the experiment in order to measure participants' normal cortisol levels. By (i) measuring salivary cortisol repeatedly during the stress response in the experiment, (ii) eliciting baseline levels of cortisol on the same time of the day outside the lab setting, and (iii) randomizing the order of treatment conditions, our experiment is designed to cleanly identify treatments effects *net of* the diurnal cycle and deviations from normal cortisol concentration due to other confounding stressors. Individuals' self-reported stress is elicited by asking subjects whether they assess the situation as threatening or as challenging and by asking them how stressed, calm, and nervous they felt before and after the work phase in each payment condition.

The two stress measures—self-reports and cortisol concentration—mark different dimensions of a stress response. Cortisol is directly related to bodily functioning and health effects—e.g. regulating energy, blood sugar and inflammation—and self-reports gauge feelings of consciously experienced stress. We elicit these separate measures for two reasons. First, we identify the effects of different incentive mechanisms on stress in both the HPA-axis and one's self-assessed experience. The one does not imply the other and by investigating both pathways we shed light on the nature of the stress response to incentives.[3] Second, on the individual level, we are

[3] An individual may undergo a peak in cortisol levels yet at the same time be unaware of such stress reaction or unwilling to recognize it. Alternatively, self-reported stress may be under- or overstated, or subject to other forms of reporting behavior that are unrelated to the true stress effect. As such, a stress reaction may be evidenced in one measure, but not in the other.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

able to document the alignment of cortisol and self-reported stress responses. This alignment, or lack thereof, is insightful as it reveals to what extent perceptions of stress can be used as a signal—both to the individual and policy maker—of HPA-axis reactivity.

The results show that tournament incentives cause stress. Salivary cortisol concentrations increase (by 57 mnol/L) during the condition with tournament incentives, but continue to decrease, i.e. follow the regular diurnal cycle, during the fixed payment condition. Increases in cumulative cortisol concentrations during the 50-min time period from the onset of a stressor to re-establishment of homeostasis are significantly higher in the tournament conditions than in the fixed payment conditions. Participants also report greater perceived stress both before and during the tournament work phase, but not in the fixed payment condition. Moreover, they appraise the tournament as more threatening and challenging. Stress responses in the private disclosure and the private disclosure tournament condition, measured either by cortisol or self-reports, are equally strong.

Notably, we find that self-reported stress is significantly correlated with our physiological stress measure. Individuals with stronger cortisol responses report higher levels of perceived stress. Partial correlation coefficients range from 0.15 to 0.28 depending on the metric that is used to measure the hormonal stress response. These results indicate that individuals' perceived stress response is aligned, albeit imprecisely, with their cortisol response. This begs the question whether individual differences in stress responses affect participants' decision to select into a tournament when having the choice to opt for a fixed payment scheme alternatively. Our findings indicate that neither objectively measured stress nor subjectively assessed stress affects the sorting decision.[4]

A final important finding concerns the design of experiments in general and the methodology of measuring cortisol responses in a lab setting in particular: we observe that participants have elevated levels of cortisol at the start of the experiment suggesting that anticipation of novel experiences in the laboratory induces a stress response. This "novelty" effect could potentially invalidate conclusions based on measures and behaviors observed at the start of laboratory experiments. An acclimatization phase could mitigate the consequences of the "novelty" effect.

Our study is complementary to the small set of studies on competitive behavior and stress. These studies compare differences in stress levels between piece rate schemes and tournament, i.e. two pay-for-performance schemes that differ by their riskiness and the scope for social comparison. The first published study in this context is by Buser et al. (2017). In their lab experiment, subjects work under a piece rate regime, subsequently play a tournament and finally work under a self-selected scheme with short delays in between. Subject provide one saliva sample in each payment condition. Buser et al. (2017) find that both cortisol levels and self-reports only significantly increase relative to baseline after the tournament

[4] Clearly, we can only study sorting decisions in our lab environment, in which acute stress responses are induced that are certainly not detrimental to health, but we cannot study whether workers would respond to chronic stress experience by avoiding tournaments.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

and not the piece rate. Buckert et al. (2017) use a within-subject design to compare stress reactions of a treatment group—with an uninterrupted sequence of work and payment phases as in Buser et al. (2017)—and a control group—where subjects receive a piece rate for each successive work phase. They find that only self-reported mood differs between the treatment and control group. In both studies, the time between work phases is short relative to the time it takes for cortisol levels to fall to reach homeostasis again (approximately 50 min). Therefore, both studies are unlikely to separate stress responses to tournament conditions form stress responses induced by piece rate schemes or merely participation in an experiment. Zhong et al. (2018) uses a design that addresses this issue and lengthens the recovery time between work phases, in an experiment that is otherwise similar to the one by Buser et al. (2017). Zhong et al. (2018) measure cortisol and alpha-amylase and show that a tournament competition induces larger changes in these biomarkers than piece rate schemes. Cahlikova et al. (2020) study whether competitiveness and performance of men and women changes when they are exposed to an external stressor. Using the Trier Social Stress Test (TSST) to induce stress in the lab they find that stress negatively affects the performance of women in tournaments, but not in a piece rate setting, while the performance of men is not affected by stress in either payment scheme.

Our study differs from the above mentioned studies in several aspects. First, rather than focusing on differential stress responses in two pay-for-performance schemes, we ask whether tournament incentives induce a stronger stress response than a payment scheme that is *independent* of performance.[5] The linkage between reward and performance is often observed in working life yet it remains unknown how acute stress responses—which are precursors for detrimental health effects—differ in situations where pay is tied to performance and when it is not. Second, we study the alignment between perceived stress responses and cortisol responses. The alignment of self-reports and cortisol reaction is insightful as it is a prerequisite for informed decision-making. If an individual is able to sense stress and can gauge the costs associated with the hormonal stress response, he or she can factor in the costs of stress in terms of adverse health when sorting into stressful conditions. Someone may avoid a stressor, for example, if the expected costs outweigh the benefits. Moreover, policy makers and researchers may utilize self-assessed stress to approximate cortisol reactivity when measure for the latter are too costly to obtain. Third, we identify different mechanisms that may induce incentive-related stress. Dickerson and Kemeny (2004) point towards uncertainty and self-evaluative threat as important drivers of stress reactions. As such, the inclusion of a treatment where the tournament's outcome is publicly announced shines light on how different incentive features may cause stress. Fourth, the participants in our experiment are all men.

[5] A recent study by Allan et al. (2021) documents the effects of a piece rate versus a minimum performance treatment. In the former, subjects receive a piece rate for each unit of performance whereas in the latter they receive a fixed amount if they are able to attain a relatively low performance threshold. Allan et al. (2021) show that stronger incentives generate greater cortisol reactions and self-reported stress than weaker incentives.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Our prime focus is on the hormonal and perceived stress responses caused by tournaments and the mechanisms that may induce incentive-related stress. By having a male-only sample we can ensure that the effect of incentives are not confounded by gender competition. Furthermore, noise in cortisol measurements is likely larger for women than for men as cortisol responses vary over the menstrual cycle and are affected by intake of contraceptives. Finally, we implement a novel experimental design that allows us to identify the causal effect of pay-for-performance on stress independent of diurnal, novelty and order effects.

The remainder of this paper is structured as follows. Section 2 describes the research design. Section 3 presents the results. Section 4 concludes the paper.

2 Design of experiment

The lab experiment is designed to elicit participants' stress responses when faced with tournament incentives as compared to a fixed payment as remuneration for performance on a real effort task. The task entails multiplications of a one-digit and a two-digit numbers and is taken from Dohmen and Falk (2011). As the salivary sampling of baseline levels of cortisol is crucial for our experiment, it is conducted in two sessions that are scheduled in two consecutive weeks.

2.1 Week 1

In the first week, we elicit demographic information as well as individuals' attitudes towards risk, losses and uncertainty, and their cognitive ability through the Raven IQ test. Furthermore, we familiarize the participants with the real effort task by having them solve the same type of multiplication problems, i.e. multiplying a one-digit number and a two-digit number, as they face in the real effort task in week two. In week 1, they work on such multiplication problems for a period of 5 min and receive a piece rate for each correct answer. Finally, we measure participants' self-assessment with respect to this task. All these measures are financially incentivized. At the end of the session, we inform participants that saliva samples will be taken in the second week and that they have to take two saliva samples at home on two different weekdays before the second session. We instruct participants on how to take saliva samples, by oral instructions, an instruction video and written instructions, in which participants are informed about rules that they have to adhere to when taking saliva samples, so that we can account for or rule out confounding factors in cortisol measurement. These rules also constrain the following activities within a certain period before taking a saliva sample: alcohol intake, exercise and sports, dentist visit, food intake, caffeine intake, smoking, teeth brushing, water intake (see the online appendix for details). As cortisol responses also vary over the menstrual cycle and are affected by intake of contraceptives, we only invite men to our experiment.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

2.2 Week 2

The second experiment session consists of three blocks, in each of which participants works on the same real effort task as in week 1. In the first two blocks, participants are once remunerated by a fixed payment, and once according to their performance in a two-person tournament. In the third block, they can choose the fixed payment or the tournament incentives for remuneration.

2.2.1 Incentive schemes

In the fixed payment scheme, subjects receive 600 points irrespective of the number of correctly solved multiplication problems. In the tournament scheme, participants are matched in pairs and play a tournament against each other. The person who solves more problems earns 2300 points, while the loser earns 0 points.[6] All participants face both incentive schemes sequentially. The order is randomized across the different experiment sessions. After having performed the task under both incentive schemes, participants are asked how they want to be rewarded for performing the effort task a third time, by choosing one of the two aforementioned payment schemes.[7] Subjects are informed about their performance under *all* incentive schemes during and immediately after the task. During the multiplication task, subjects observe in real-time how many multiplications they have solved correctly so far. At the end of each multiplication task, subjects see their total correct score displayed on a separate screen. The outcome of the tournament(s) is announced at the end of the experiment when all tournaments have been played. Disclosure of the tournament outcome takes place at the end to minimize income effects on stress and decision-making.

In addition, we investigate whether a public announcement of each participant's outcome in the tournament elicits a greater stress response than a private announcement. The self-preservation theory posits that "socio-evaluative threat occurs when an important aspect of the self-identity is or could be negatively judged by others. We propose that social-evaluative threat is most likely to occur when failure or poor performance could reveal lack of a valued trait or ability" (Dickerson & Kemeny, 2004). Consequently, to test whether subjects show greater stress responses in face of such socio-evaluative threat, subjects either reveal their outcome to the other participants or keep the information to themselves during the experiment. In particular, in the *private* treatment subjects remain seated after receiving information about the tournament's outcome. Alternatively, in the *public* treatment, subjects are asked to

[6] If both participants solve the same number of problems, the computer randomly selects the person to earn 2300 points. If an odd number of participants participate in the session, one randomly chosen participant's output is used a second time and compared to the output of the unmatched participant so as to determine the earnings of the unmatched participant.

[7] Subjects were informed that the tournament would not take place if only one participant in a session enters the tournament, and that all participants would be paid according to the fixed payment. In every session, however, at least two subjects selected the tournament payment so that tournaments took place in all sessions. Therefore, all subjects could be remunerated according to their chosen payment scheme.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

step outside of the cubical after have being informed about the tournament's result. Subsequently, subjects receive a rolled-up paper (thereby keeping its inscription hidden) that stated, "WINNER", "LOSER", or "UNDECIDED" depending on their outcome in the tournament. Finally, after having formed a circle, subjects are asked one by one to show the writing on their outcome-sign. This procedure is communicated to the participants at the same time when the tournament scheme is explained.

Whilst all subjects solve multiplications both under the tournament and fixed payment scheme, subjects undergo either the private or public tournament treatment. Hence, to assess the effect of the private versus the public treatment on stress, we look at between-subject differences. When analyzing the effect of the tournament versus the fixed payment on stress, we make both between- and within-subject comparisons, however.

2.2.2 Measurement of salivary cortisol

During the experiment, we repeatedly measure salivary cortisol to track participants' stress response along the hypothalamic-pituitary-adrenocortical (HPA) axis. The HPA-axis is activated in response to psychological stressors.[8] As part of the physiological stress response, cortisol—a steroid hormone that is produced in the adrenal cortex—is released into the bloodstream and becomes measurable in saliva with a short delay.[9] The release of cortisol into the bloodstream takes about 5–20 min; and free cortisol in the blood is transferred to saliva within no more than 2–3 min (Bozovic et al., 2013). Thereafter cortisol decays and homeostasis is reached after 40–50 min after the onset of the stress reaction. To track this cortisol response, we take saliva samples at the onset of each work phase and 20, 30 and 50 min after the start of the payment condition.[10] Thereby we capture baseline cortisol, the peak response and the return to baseline.[11] The exact moments are denoted in Figure A1 in Appendix A.

When measuring stress levels through cortisol there are several aspects that need to be taken into consideration. First of all, cortisol follows a diurnal cycle, where cortisol levels are higher upon waking with peaks after 30 min (the so-called cortisol awakening response), and steadily decrease from this peak throughout the rest of the day. In particular, on average, cortisol levels show a small and stable decrease

[8] The HPA axis is one of the major neuroendocrine systems and comprises complex interactions between the hypothalamus, the pituitary gland and the adrenal glands. It regulates digestion, energy storage and expenditure, mood and emotions as well as the immune system, and "is vital for supporting normal physiological functions and regulating other systems'" (Dickerson and Kemeny, 2004, p. 356).

[9] The stress reaction along the HPA axis starts with the secretion of corticotropin-releasing hormone (CTH) by the hypothalamus. Subsequently, the adrenocorticotropic hormone (ACTH) is released by the anterior pituitary gland, which leads to the release of cortisol, a steroid hormone that is produced in the adrenal cortex, into the bloodstream.

[10] Saliva samples are well suited for the purpose of cortisol elicitation, since the secretion of free cortisol into blood, i.e., biologically active cortisol, maps well into saliva (Kirschbaum and Hellhammer, 1994).

[11] Participants are exposed to a new payment condition only 50 min after the onset of the previous payment condition, in order to ensure that homeostasis is reached when the new payment condition begins.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

in the hours after 2 pm (e.g. Molitch, 1995; Weitzman et al., 1971).[12] We therefore ran all experimental sessions at the same time of the day and started all sessions at 2 pm. Any changes in cortisol levels that are caused by the diurnal cycle are therefore approximately equal across sessions. To increase accuracy, and possibly filter out extreme cases, we administered a short questionnaire at the end of the experiment asking subjects about their wake-up times and health.

Furthermore, individual stress responses should be measured as deviations from individual-specific baseline levels as both baseline levels and stress responses vary substantially between individuals. Baseline levels are measured in two ways: (1) pre-treatment cortisol levels during the second experiment and; (2) cortisol levels that are elicited at home on two working days prior to the second experiment. We incentivized all participants to collect two cortisol measurements at home at the same time as the pre-treatment baseline measures in the lab, i.e. 2.30 pm. The home samples reveal individuals' regular cortisol levels, i.e. under conditions outside the lab. In particular, home samples are arguably not confounded by anticipation effects that are related to the experiment (e.g. individuals might be nervous at the onset of the experiment). In order for participants to know how to collect salivary samples at home, we implemented a design with two sessions in two consecutive weeks, and instructed participants how to collect the saliva samples in the first week. In the first week they received instructions on saliva sampling, storage of saliva samples in freezers at home and transportation to the lab in the second week.[13] To gain assurance that subjects produce the saliva swaps at the right time of day, they are asked to hand in a selfie that shows their face, the respective saliva swab, and the time and date.[14] Subjects received 4 (10) euro if they delivered one (two) home sample(s) of saliva and could document that it was properly elicited.[15]

2.2.3 Quantification of acute cortisol response

To quantify and evaluate a cortisol response, i.e. to map each cortisol level per payment scheme into a measure of a cortisol response, we follow Pruessner et al. (2003) and Miller et al. (2013). Pruessner et al. (2003) propose the area under the curve with respect to the increase (AUCI) as it "simplifies the statistical analysis and increases the power of the testing without sacrificing the information contained in multiple measurements".[16] The formula, for a given individual in treatment block X,

[12] Typically, studies provide evidence of the diurnal cycle by asking participants to measure salivary cortisol levels during a typical day, i.e. involving standard activities for that individual (Adam et al., 2006; Cohen et al., 2006; Smyth et al., 1997; Stone et al., 2001). In particular, Stone et al. (2001) document an approximately linear afternoon decrease for male German college students.

[13] Details on the exact experimental procedures can be found in the online appendix.

[14] By typing "time" into the google search bar, the current time and date are displayed on the computer screen.

[15] The Institutional Review Board (IRB) of Maastricht University (ERCIC) approved the research protocol outline above in May 2016 with reference ERCIC_010_02_04_2016. This approval decision was taken over by the University of Bonn.

[16] With 1721 citations on www.google.scholar.com (at the time of the experiment), this article features the most prominently used cortisol response estimates for *acute* stress responses that involve several cortisol observations per stressor.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

is given by $\frac{(C_{X1}+C_{X2})\cdot 20}{2} + \frac{(C_{X2}+C_{X3})\cdot 10}{2} + \frac{(C_{X3}+C_{X4})\cdot 20}{2} - C_{X1} \cdot 50$ where C_{X1} to C_{X4} correspond to the four cortisol levels measured at the onset, 20th, 30th and 50th min of treatment X. AUCI measures the change in cortisol concentration over time relative to the initial cortisol measurement. The formula is simply the total cortisol concentration during treatment block X minus the surface below the horizontal of C_{X1}. Thereby the initial value is interpreted as the baseline or 'normal' value to which a response should approximately return after an event has occurred.

Miller et al. (2013) develop a cortisol response threshold to distinguish responders from non-responders. Subjects who experience a percentage baseline-to-peak increase of 15.5% or greater are classified as responders.[17] Subjects with a lower response are marked as non-responders. To assess the baseline-to-peak increase we measure the percentage change between the first cortisol measurement of the treatment (i.e. C_{X1}) and the maximum of the second and third cortisol measurement (i.e. C_{X2} and C_{X3}). Next to the comparison of AUCIs that show a *relative* treatment effect, the threshold determined by Miller et al. (2013) is able to determine whether a change in cortisol levels is large enough to constitute an effect to the stressor.

2.2.4 Subjective stress

Since we are interested in whether participants perceive the stressfulness of fixed payment incentives and tournament incentives differently, we measured participants' appraisals and perceived stress at multiple instances during the experiment. This also allows us to analyze whether self-reported stress and cortisol responses are aligned—i.e. whether individuals with a greater cortisol response also report higher levels of perceived stress? We elicited subjective measures of stress by asking participants to rate how stressed, calm, and nervous, they feel just before they had to work on the real effort task in the tournament condition and in the fixed payment condition. Participants stated how their agreement with the statement "Right now, I feel calm/nervous" on a 5-point Likert scale that ranges from "not at all (1)" to "very much (5)". Also, directly after the task, we asked participants on the same scale how stressed they were during the task.

We also assessed individual's primary appraisals of the potential stressors. Primary appraisals indicate whether an individual appraises an event as irrelevant, positive or dangerous, i.e. primary appraisals capture the impact of a situation on one's well-being. Stress is hypothesized as being the result of a cognitive appraisal process where individuals appraise the situation as potentially dangerous.[18] We use a questionnaire (PASA) designed by Gaab et al. (2005) to capture primary appraisals.[19] A primary appraisal that causes stress consists of three dimensions: harm/loss, threat,

[17] Cortisol is secreted in a pulsatile fashion and can therefore increase and decrease solely due to the discharge mechanism. Hence, a cortisol increase must be large enough to constitute a true treatment effect.

[18] Subsequently, these appraisals result in an emotional, physiological, and behavioral stress response (Folkman, 1984; Gaab et al., 2005).

[19] The PASA questionnaire also elicits secondary appraisals which relate to "the evaluation of coping resources and options" (Folkman, 1984). This appraisal evaluates which coping resources, including social, psychological, physical and material assets, are applicable to dealing with an event. We focus

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

and challenge. Harm/loss refers to damage that has already occurred. Threat indicates the potential of harm or loss, and challenge relates to the opportunity for mastery, growth or gain (Folkman, 1984). The questionnaire includes items about the threat and challenge appraisals. In line with Gaab et al. (2005), we disregard harm\loss appraisals "…as we set out to operationalize anticipatory stress processes, […] and not the appraisal of past stressful events". The participants indicate the strength of their appraisal on a 6-point Likert scale that ranges from "strongly disagree (1)" to "strongly agree (6)".

2.2.5 Sequence of events

During the experiment, subjects perform the real effort task in three *blocks* where they receive either a fixed payment or tournament payment. Each block lasts approximately 50 min and consists of different *phases* in which different stress measurements are elicited. The three phases are setup as follows:

a. *Anticipation phase (approximately 5 min)*

At the beginning of the first phase, minute 0, participants are requested to provide a saliva sample. Participants have not received any information about the subsequent treatment yet, and hence this sample serves as a baseline value of cortisol level at the start of the experiment. Subsequently, participants are informed that they will perform the multiplication task for a period of 10 min and will be rewarded according to the incentive scheme they are assigned to. Shortly after this announcement and just before starting the work phase, we assess to which extent participants perceive working under the particular incentive scheme as stressful, challenging and/or threatening, using the multiple subjective items mentioned above. Each subject thus indicates how he feels in anticipation of the task with different incentive schemes.

b. *Work phase (approximately 11 min)*

During the second phase, participants perform the real effort task for a period of 10 min. Directly afterwards, to gain a measure of participants' perceived stress levels, we ask participants to indicate how stressed, nervous, and calm they felt during the differently incentivized multiplication tasks. In particular, the retrospective items state: "How stressed did you feel during the previous 10 min?" and "During the

Footnote 19 (continued)
solely on primary appraisals, as secondary appraisals are less relevant for this study. We decided not to include the secondary appraisals from PASA as the theoretical construct and questionnaire items do not match with the relatively simple task and could be confusing to participants. For example, items include "In this situation I can think of lots of action alternatives."

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

previous 10 min I felt calm/nervous". Participants can answer on a 5-point Likert scale that ranges from "not at all (1)" to "very much (5)".

c. *Recovery phase (approximately 34 min)*

Finally, we allow subjects´ cortisol responses to recover from the treatment such that their stress levels return to baseline when they enter the next block. Therefore, after participants have finished the multiplication task and stated their feelings about stress, they are asked to fill in questionnaires for the remaining 34 min. Upon finishing the questionnaire, they were allowed to read magazines that they selected upon arrival at the lab.[20] These remaining 34 min are necessary to return to homeostasis.

To capture the peak in cortisol concentrations, we take saliva samples in the 20th and 30th min of the first and second block. Since cortisol is expected to return to its baseline value within 50 min after the announcement of the stressor, participants are again required to provide a saliva sample at the 50th min during the first and second block. Notice that the cortisol measurement at the 50th min of the first and second block also serve as the measurement at the 0th min of the subsequent block. In the last block, we did not take a saliva sample in the 50th min as is no subsequent treatment. In total, we therefore elicit eight saliva samples during the second experiment session. Immediately after the last saliva elicitation, participants are informed about the tournament's outcome. Depending on the respective treatment, subjects either reveal their outcome to the other subjects or remain seated in the cubical. Finally, participants receive information about their total payoff and are called forward one-by-one to collect their payoff. In Appendix A, Figure A1 depicts the timeline of the experiment in week 2 for participants who start with the fixed payment in block 1 and subsequently play the tournament in block 2.

[20] Magazines included National Geographic, The Economist, and Bloomberg Businessweek. The articles were chosen as to cause as little stress as possible. With respect to the questionnaires in the first block, we measure participants' optimism with the Revised-Life Orientation Test (Scheier, Carver, and Bridges, 1994). Subsequently, a measure of participants' locus of control is elicited by using the locus of control items from Nolte et al. (1997), which is based on Rotter's (1966) original construct. Next, we used the Revised Competitiveness Index (Houston et al., 2002) to measure interpersonal competitiveness. We then proceed with the Big Five personality test (Costa and McCrae, 1992), and end this block by measuring participants' perception of their masculinity by asking participants to fill in the masculinity items from Persönlichkeit und Gesundheit (SOEP pretest questionnaire, in the field 2007). In the second block, participants' level of self-efficacy is elicited by adopting the New General Self Efficacy Scale (Chen et al., 2001). Thereafter, we proceed by measuring fear, shame, contempt, and shyness by asking the respective questions from the fourth edition of the Differential Emotions Scale (Izard, 1993). Also, to describe participants' individual differences in interpretation of events, and the relation of this interpretation to their thinking and behavior, we use the attributional complexity scale (Fletcher et al., 1986). Finally, at the end of the third block, we gain information about participants' demographic characteristics, and health and lifestyle indicators that potentially influence cortisol responses. Statistical associations between aforementioned questionnaire items and stress responses are not documented in this paper and are available upon request.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

2.3 Procedural details

We conducted six sessions at two European universities—three in Maastricht and three in Bonn—during the summer of 2016. Each session involved two visits to the lab in two consecutive weeks. In total 99 men completed the experiment, 34 in Maastricht and 65 in Bonn (see Table A1 in Appendix A for details).[21] During three sessions the outcome of the tournament was announced privately, whereas in the other three sessions the outcome was announced publicly. With respect to the order of the first two blocks, in which participants were subjected to the fixed and tournament payment, 4 sessions initiated with the fixed payment (FT) and 2 sessions begun with the tournament scheme (TF). Any form of communication and use of electronic devices (calculators, phones etc.) was strictly forbidden throughout the experiment. Economics students in their third or fourth year were excluded from participating in the experiment, as they may be acquainted with principal-agent theory and primed to behave accordingly. On average, the first session lasted about 1 h, and the second session took approximately 3 h to complete. After the provision of the instructions on salivary sampling at the end of week 1, participants were asked whether they are willing to sign the informed consent statement. They were told they could stop at any time during the experiment, and were informed that if they chose not to sign the statement, they could still join the experiment, and receive payment for all other parts of the experiments.[22] The average earnings were 85 euro. The conversion rate from points to euro was 1.7 eurocents per point. The experiments were programmed in z-Tree (Fischbacher, 2007).

In our analysis, we restrict the sample to 96 men.[23] One participant's computer malfunctioned such that he could not play the pre-determined tournament during the second block. While the problem was resolved, we discard the observation since the event per se might have induced a stress response. Moreover, subsequent decision-making might have been affected. For another participant, cortisol levels could not be determined for 5 saliva samples by the laboratory in Dresden. We also discard the observation of another participant who has extreme values of cortisol. In particular,

[21] A total of 6 out of 105 subjects did not participate in week 2. We find no statistical difference in background characteristics of those who did not participate in week 2, but as the attrition rate is low, we lack power to make any meaningful statements about self-selection into second session of the experiment.

[22] All participants who participated in the experiment signed the informed consent form.

[23] To estimate the amount of subjects required to detect statistically significant treatment effects, we conduct a power analysis with the following parametrization. We test two main hypotheses: (i) a tournament causes greater stress than a fixed payment and (ii) a public tournament causes greater stress than a private tournament. The expected composite effect size of the former is 0.9 which originates from (Buckert et al., 2017 and Zhong et al., 2018). The expected effect size is likely an underestimate as the aforementioned studies compare the effects of tournament versus piece rates whereas we compare tournament versus fixed payment. The expected effect size of the second hypothesis is 0.5 which comes from Gruenewald et al. (2004). In their setup *without* financial incentives, stress is analyzed for subjects who had to perform speech and math tasks either in "the presence of an unfriendly, evaluative audience " or in private. We set the significance level to 0.10 and statistical power to 0.80. By calculating the minimum detectable effect size (Spybrook et al., 2011) we find that more than 35 subjects are needed to confirm hypothesis i) and more than 90 subjects are required to confirm hypothesis ii). To be safe, we invited a conservative total of 105 subjects to the experiment.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

both the levels *and* changes are greater than 5 times the standard deviation from the mean. In addition, the participant indicated that he had more than 10 alcoholic beverages the evening previous to the experiment and, on average, drinks more than 30 alcoholic beverages per week.

Saliva samples were collected with a Salivette® Tube (Sarstedt, Nümbrecht, Germany). The samples were stored for approximately 3 weeks at −20 degrees Celsius. Subsequently, the saliva samples were centrifuged for 3 min at 3000 rpm. Salivary cortisol levels were determined at the Dresden LabService GmbH by the Hettich Centrifuge.

3 Results

Before describing the main results of the experiment, descriptive statistics are shown in Table 1. Table 1 shows that the randomization between the public and private treatment is successful as the averages of observables are mostly indistinguishable. Moreover, the table shows that subjects solve more multiplications correctly under the mandatory tournament incentives than a fixed payment (paired t-test; $\Delta=-6.9$; $p=0.001$).[24] This finding also holds for both the private and public treatment, and the self-selected incentives. Finally, 58 percent ($N=56$) of the subjects select the tournament in the third block.

We start the analysis of the main results by assessing whether a tournament and a fixed payment scheme generate different stress responses (Sect. 3.1). In Sect. 3.2 we investigate the relation between acute cortisol responses and self-stated measures of stress. Finally, we investigate whether individuals' decision to select a tournament or a fixed payment is influenced by their stress reaction in Sect. 3.3.

3.1 The effect of incentives on stress responses

3.1.1 Acute cortisol responses

Figure 1 depicts the development of cortisol during the experiment. As described in Sect. 2, 64 subjects first worked for a fixed payment and subsequently under tournament incentives ($i \in$ FT). Their average cortisol development is depicted by the solid curve in Fig. 1. For the remaining 32 subjects the treatment order is reversed ($i \in$ TF) and the cortisol development is shown by the dashed curve in the figure.

Figure 1 describes the development of average cortisol levels for both treatments (fixed payment and tournament) and both orders of treatments (TF and FT). The solid curve shows de development for subjects who start with the fixed payment, subsequently play the tournament, and finish with the selection phase ($n=64$). The

[24] It should be noted that subjects solve less multiplications in the first block than in the second block. For the fixed payment this difference is statistically significant at the 5% level. We conjecture that there may be learning effects and that subjects feel stressed at the beginning of experiment such that performance is affected. Below, we provide evidence for increased stress levels at the onset of the experiment.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Table 1 Descriptive statistics (means and standard errors)

	Total (N = 96)	Public (N = 46)	Private (N = 50)
Age	22.32 (0.30)	22.15 (0.47)	22.48 (0.39)
Study year	2.56 (0.20)	2.39 (0.32)	2.72 (0.25)
IQ (Raven)	5.36 (0.21)	5.07 (0.28)	5.64 (0.30)
Productivity	23.2 (1.00)	22.28 (1.44)	24.04 (1.39)
Self assessment	0.62 (0.03)	0.6 (0.04)	0.64 (0.04)
Risk attitude	0.44 (0.01)	0.47 (0.02)*	0.41 (0.02)
Performance fixed payment (mandatory)	49.02 (2.19)	46.67 (2.85)	51.18 (3.28)
Performance tournament (mandatory)	55.93 (2.29)	53.83 (3.08)	57.86 (3.36)
Incentive choice (Tournament = 1)	0.58 (0.05)	0.57 (0.07)	0.6 (0.07)
Performance fixed payment (selected)	29.6 (3.39)	31.45 (4.60)	27.75 (5.07)
Performance tournament (selected)	64.73 (2.82)	62.38 (3.98)	66.77 (4.00)

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

dashed curve indicates the cortisol development of subjects with the alternate order: TF ($n = 32$). *home* indicates the average of the individual averages of the two cortisol home samples taken at approximately 14.30 earlier in the week.

The figure depicts two results. First, cortisol levels increase after the onset of the tournament condition and peak after 20–30 min independent of the order. In the fixed payment treatment, no increase is observed. On the contrary, cortisol levels decrease steadily after the onset of the fixed payment condition, consistent with the decrease in cortisol during the diurnal cycle. These differential developments of cortisol indicate that a tournament generates greater cortisol responses than a fixed payment. A between-subject comparison corroborates this finding: cortisol levels are elevated for participants working under tournament incentives relative to cortisol levels of participants working for fixed payments, independent of order. In particular, the left panel of Fig. 1 shows that 20 min after the onset of the stressor, cortisol concentration is higher among workers in the tournament condition (dashed line) than among workers in the fixed payment condition. Cortisol concentrations also rise for workers who face the tournament incentives in the second work period (solid line in the right panel of Fig. 1) from baseline (50 min) to peak (70–80 min).

Table 2 shows that tournament incentives cause greater cortisol responses than fixed payment schemes. Panel A shows mean and median comparisons between cortisol responses that are caused by the tournament (*mean* T and *med.* T) and fixed payment (*mean* F and *med.* F) during the first block. Panel B describes these estimates for subjects in the second block.[25] During the first block the AUCI of the tournament is greater than that of the fixed payment. This result is statistically significant at the 5% level using parametric (*t*-test) and nonparametric (Mann–Whitney

[25] In Table 2A we do not control for session effects. Nevertheless, if we consider the average effects of incentives on cortisol—i.e. *mean* F and *mean* T—and control for session effects in an OLS specification, then the results are the same.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

U) hypothesis testing. With respect to the second block, both the AUCI of the tournament exceeds that of the fixed payment. The difference in AUCI measures is statically significant at the 5% level (both parametrically and nonparametrically). The average baseline-to-peak increase in the tournament treatment is greater than 15.5% irrespective of the treatment order. The median baseline-to-peak increase is greater than 15.5% in the second block, but not in the first block. During the first and second block, 41% and 53% of the subjects that play the tournament are classified as cortisol responder, respectively.

Significant differences between median responses indicate that results are not driven by a few outliers. Moreover, differences between mean and median effects suggest that stress responses are skewed. Nevertheless, the effects of incentives on stress are not driven solely by an increase in skewness. Figure A2 in Appendix A shows that in both orders, i.e. independent of whether subjects first faced the tournament incentive or the fixed payment, the distribution of cortisol concentrations is shifted to the right in comparison to the distribution of cortisol concentrations in the fixed payment scheme.

The second striking feature is that cortisol levels are elevated during the first measurements. This suggests a *novelty effect*—i.e. a positive cortisol response due to a new or unanticipated experience.[26] Four observations indicate that the initial increase is in fact due to the participation in the experiment. First, salivary cortisol before exposure to the tournament is higher for subjects who are first exposed to tournament incentive (i.e. subjects in sequence TF) in comparison to subjects who play the tournament in the second block, while the absolute increase of cortisol after exposure to the tournament is smaller for these subjects [$\Delta_{TF,2-1}-\Delta_{FT,5-4}=-0.65$ (0.56, $p=0.25$)]. Second, and consistent with the first finding, the subsequent decrease in cortisol after the peak caused by tournament incentive is larger for subjects in TF [$\Delta_{FT,7-5}-\Delta_{TF,4-2}=1.64$ (0.35, $p=0.00$)]. Third, for subjects who start with the fixed payment treatment, i.e. in sequence FT, we observe a stronger decline during the fixed payment phase than for subjects who are on fixed payment incentives in the second block [$\Delta_{FT,4-1}$-$\Delta_{TF,7-4}=-1.95$ (0.53, $p=0.00$)]. The greater decrease in cortisol levels under the fixed payment scheme during the first block indicates that subjects have elevated cortisol levels due to stress at the beginning of the experiment which subsequently recover during the first block. This also holds for the tournament in the first block. Fourth, the individual averages of the two cortisol home samples (c_h)—which are collected on earlier days at the same time that the experiment begins—are smaller than the first cortisol measurement during the experiment [$c_1 - c_h=2.27$ (0.50, $p=0.00$)]. In comparison to cortisol home levels, cortisol levels increase by 89% on average in the baseline measurement. Approximately 58% of the subjects experience an increase that is greater than 15.5% which

[26] The effect may originate from alternative sources than novelty. By *novelty effect* we mean the initial cortisol response that is caused by participating in the experiment independent of its content. Kirschbaum et al. (1995) have reported cortisol responses when subjects were purposefully faced with an unanticipated situation. We argue that such effect may also take place at the beginning of a laboratory experiment.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

provides further evidence that participation in the experiment induces a cortisol response.[27]

These observations indicate that subjects' cortisol levels under normal circumstances are lower than the ones observed at the beginning of the experiment. As cortisol reactivity in the first block is potentially affected by the novelty effect, within-subject testing of cortisol responses to incentives must be interpreted with caution.[28] Table A2 in Appendix A shows effects of incentives on cortisol responses in a within-subject comparison. Moreover, in the Appendix B we also disentangle the multiple latent effects at work: treatment (i.e. fixed payment and tournament), novelty and diurnal effects. Apart from the impact of the novelty effect on the measurement of cortisol responses to incentives, the effect has important implications in general for the design of laboratory experiments and potentially the collection of survey data. As stress potentially affects the quality of economic decision-making (Mani et al., 2013), we should take into account that decision-making may be affected by merely participating in an experiment or survey. Moreover, the novelty effect may be correlated to behaviors that are elicited in the experiment or survey. Consequently, estimated relationships may be biased. We discuss this possibility with respect to sorting into incentive schemes below.

3.1.2 Subjective stress

In this section, we investigate the association between incentives schemes and subjective stress—i.e. self-stated feelings of stress and primary appraisals. The first measure of subjects' feelings of stress indicates how stressed, nervous, and calm they felt during the differently incentivized multiplication tasks. The three measures are highly correlated (approximately 0.7) and therefore we take the average to reduce measurement error.[29] The second measure states how calm and nervous they felt right before performing the multiplication tasks. Again, given the high correlation between measures (approximately 0.6) we take the average to reduce measurement error. The upper and lower panel of Table 3 show the results for orders FT and TF, respectively, where we consider within-subject differences. In both treatment orders, subjects indicate that they are significantly more stressed in the tournament scheme. Parametric and non-parametric tests reveal that these results are statistically significant (t-test: p-value < 0.001; Mann–Whitney test: p-value < 0.001).[30] In

[27] We acknowledge that salivary cortisol concentrations in home samples may be elevated relative to baseline if subjects experienced stressors at home prior to saliva sampling. If this is the case, i.e. if the cortisol levels of the home samples are elevated, then the estimate of the novelty effect is a *lower bound* and the actual effect is larger in comparison to cortisol baselines in the lab.

[28] As indicated in Dickerson and Kemeny (2004), most cortisol responses return to baseline within 50 min. Hence, we do not expect the novelty effect to transfer to the second block cortisol measurements.

[29] The item response for calmness is reversed, such that greater values indicate lower feelings of calmness. Moreover, principal component analysis of the three variables suggests that the first component explains approximately 80% of the total variation. Moreover, the absolute weights of the eigenvector are approximately equal serving the interpretation that the first principal component captures an overall measure of stress. Taking the average is therefore warranted.

[30] A between-subject comparison yields the same quantitative results.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Table 2 Cortisol responses during block 1 and 2 (between-subject)

	mean T	*mean* F	*median* T	*median* F
A. 1st block ($n_T=32$, $n_F=64$)				
AUCI	−17.4**	−80.6	−16.2*	−60.9
baseline-to-peak (%)	37.5%*	−4%	4%***	−23%
B. 2nd block ($n_T=64$, $n_F=32$)				
AUCI	29.5***	−37.2	10.1***	−24.3
Baseline-to-peak (%)	49% ***	−9%	19%***	−14%

Depicts the mean and median of cortisol responses that occurred in the first (A.) and second (B.) block. With respect to the means, statistical significance is tested with an unpaired two-sided *t*-test of the difference of the means. For the median a two-sided Mann–Whitney *U* test of the difference of the distributions is implemented. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

addition, Table A3 in the Appendix A shows that participants appraise the tournament both as significantly more threatening and challenging.

3.1.3 Private and public disclosure of tournament's outcome

Socio-evaluative threat is hypothesized to be a factor that affects acute cortisol responses (Dickerson & Kemeny, 2004). To assess the impact of this factor, we have designed the experiment such that the tournament's outcome is either publicly or privately announced. In the public treatment, subjects reveal the outcome of the tournament to the other subjects in turn. That is, they receive a sign that states "WINNER", "LOSER", or "UNDECIDED" and show it to the other participants. In the private context, participants remain in their cubical and no announcement takes place.

First, Table 4 reveals the acute cortisol responses for the public (PU) and private (PR) settings. Again, we use AUCI and baseline-to-peak response to estimate cortisol responses to the tournament treatments. In the upper and lower panel of Table 4, the results are depicted for subjects who played the tournament in the first and second block, respectively. Neither AUCIs nor baseline-to-peak increases are significantly different between the socio-evaluative treatments.

Second, Table 5 depicts the effect of public and private tournament announcement on the self-stated feelings of stress before and during the multiplication task. We observe no statistically significant differences between the private and public treatment. In addition, subjects appraise the tournaments types as equally threatening and challenging (see Table A4 in Appendix A).

3.2 Acute stress responses and self-stated stress

We have shown that tournament incentives increase *both* acute cortisol responses *and* self-stated measures of stress. Next, we examine whether these measures are related on the individual level. Do individuals who state they experience greater stress also have greater cortisol responses? This question is important for multiple

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

reasons. First, stressful situations can generate health problems through prolonged cortisol increases. For this reason, individuals may or may not shy away from a situation depending on the extent to which they perceive it as stressful. If the perception of stress, however, does not correspond with the underlying health cost—i.e. cortisol responses—then this may lead to detrimental outcomes. Second, the alignment between self-stated stress and cortisol responses has shown mixed-results in the literature (Campbell & Ehlert, 2012). We contribute by showing (partial) correlations between multiple self-stated stress measures and cortisol responses.

Table 6 shows the coefficients of the correlation between cortisol responses—i.e. AUCI and baseline-to-peak increases (%)—and self-stated stress measures—i.e. perceived stress before and during the task and primary appraisals. We depict coefficients for each treatment-period cluster: tournament in block 1 (T1, $n=32$) and block 2 (T2, $n=64$) and fixed payment in block 1 (F1, $n=64$) and block 2 (F2, $n=32$). The first two rows show consistently positive correlation coefficients between cortisol responses and self-stated stress. In each treatment-period cluster greater perceptions of stress are associated with greater responses in cortisol. The last two rows depict correlation coefficients for primary appraisals. For both challenge and threat appraisals the coefficients are concentrated around zero with a minimum and maximum of -0.30 and 0.26, respectively.

The separate estimates of the correlations between self-stated stress and cortisol responses suggest a positive association: Individuals who perceive the task as more stressful generate greater cortisol increases in response to the task. To increase statistical power and account for dependent sampling, we pool the sample and estimate a random-effects regression model. In addition to the linear dependence between self-reported stress and cortisol responses, the model accounts both for treatment- and period effects.[31] The model shows how much variation in cortisol responses is explained by self-stated measures for a given treatment and period. Moreover, by observing the change in the treatment effect on cortisol after including self-stated stress as a predictor, we can observe to which extent the treatment effect is explained by self-reported stress. Notice the regression coefficients of self-reported stress should not be interpreted as causal effects. We merely represent the underlying relations between cortisol response and self-stated stress and to which extent the causal treatment effect on cortisol can be explained by individual differences in self-reported stress.

Table 7 represents the results of the random-effects regression in which AUCI is the dependent variable. The first column (i.e. baseline) replicates the results of Table 2. The second and third column show that self-stated stress before and during the task relate positively and significantly to AUCI, respectively. Keeping both the period and treatment fixed, a one standard deviation increase in self-reported stress before the task is associated with an AUCI increase of 18.23 mnol/L. The partial

[31] In particular, for subject i in period t (block 1 or 2) we estimate.

$$y_{it} = \alpha + \beta_S S_{it} + \beta_t t_{it} + \beta_T T_{it} + \beta_I t_{it} T_{it} + \mu_i + \epsilon_{it}$$

where y_{it} is either AUCI or $\ln(\frac{c_{max}}{c_{start}})$, S_{it} is self-stated stress, t_{it} denotes the period, T_{it} is the treatment. μ_i and ϵ_{it} are independent.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

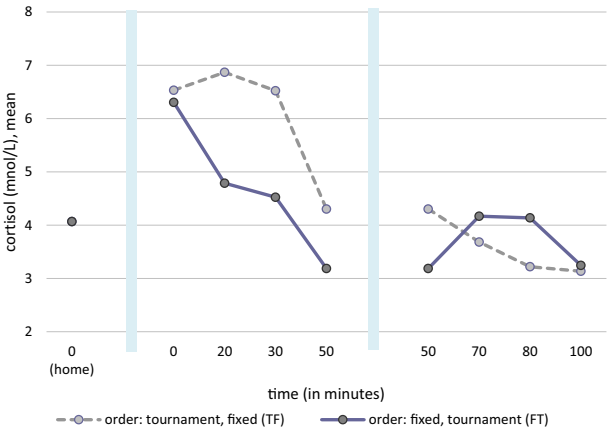


Fig. 1 Development of average cortisol levels for the fixed payment and tournament treatment

Table 3 Self-stated stress before and during task

	mean T	*mean* F	*median* T	*median* F
A. SUBJECTS IN FT (n = 64)				
Stress during task (retrosp.)	3.72***	2.40	4***	2.16
Stress before task	3.51***	2.03	3.5***	2
B. SUBJECTS IN TF (n = 32)				
Stress during task (retrosp.)	3.61***	1.98	3.67***	2
Stress before task	3.18***	1.86	3.5***	1.5

Denotes the mean and median of self-stated stress during the task in incentive orders FT (A.) and TF (B.) With respect to the means, statistical significance is tested with a paired two-sided *t*-test of the difference of the means. For the median a two-sided Mann–Whitney *U* test of the difference of the distributions is implemented. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

correlation coefficient is 0.15 ($p = 0.05$). A one standard deviation increase in self-stated stress during the task increases AUCI by 23.89 mnol/L. In this case the partial correlation coefficient equals 0.19 (p = 0.01). The results reinforce the positive associations shown in Table 6.

In addition, the self-reported stress measure partially explains variation in cortisol that is caused by the treatment effect. For both self-reported measures in column 2 and 3, the tournament effect decreases in comparison to the baseline model in column 1. Self-stated stress before the treatment is able to explain roughly 35% of the treatment effect on AUCI.[32] Approximately 45% of the treatment effect is runs through

[32] This follows by taking the difference between the total treatment effect—i.e. baseline model – and the adjusted treatment effect – i.e. model 1. The portion of the treatment effect that is then explained by self-stated stress before the task equals $1 - \left(\frac{45.76}{63.22}\right) = 0.28$ for period 1. For period 2, the ratio is equal to $1 - \left(\frac{41.05}{66.73}\right) = 0.38$.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Table 4 Cortisol responses to private and public disclosure

	mean PR	*mean* PU	*median* PR	*median* PU
A. 1st block ($n_{PR}=21$, $n_{PU}=11$)				
AUCI	−33.6	13.6	−60.8	−6.4
baseline-to-peak (%)	39%	36%	0.5%	9%
B. 2nd block ($n_{PR}=29$, $n_{PU}=35$)				
AUCI	38.2	22.4	9.4	12.3
baseline-to-peak (%)	51%	48%	12%	20%

Denotes the mean and median of cortisol responses by the private (PR) and public (PU) treatment that occurred in the first (A.) and second (B.) block. With respect to the means, statistical significance is tested with a paired two-sided *t*-test of the difference of the means. For the median a two-sided Mann–Whitney U test of the difference of the distributions is implemented. ***$p<0.01$, **$p<0.05$, *$p<0.1$

Table 5 Self-stated stress responses to public and private disclosure

	Mean PR	*Mean* PU	*Median* PR	*Median* PU
Stress before task ($n_{PR}=50$, $n_{PU}=46$)	3.53	3.27	3.5	3.5
Stress during task ($n_{PR}=50$, $n_{PU}=46$)	3.69	3.68	4	4

Denotes the mean and median of self-stated feelings of stress before (A.) and during (B.) the task by the private (PR) and public (PU) treatment. With respect to the means, statistical significance is tested with a paired two-sided *t*-test of the difference of the means. For the median a two-sided Mann–Whitney *U* test of the difference of the distributions is implemented. ***$p<0.01$, **$p<0.05$, *$p<0.1$

Table 6 Coefficients of correlation between cortisol responses and self-stated stress

	AUCI				*Baseline-to-peak Δ(%)*			
	T1	T2	F1	F2	T1	T2	F1	F2
Stress before task	0.17	0.10	0.22*	0.03	0.31*	0.06	0.22*	−0.08
Stress during task	0.25	0.22*	0.2	0.06	0.34*	0.19	0.27**	0.29
Challenge appraisal	0.08	0.12	−0.03	−0.13	0.26	0.08	0.19	−0.03
Threat appraisal	−0.13	0.01	0.14	0.09	0.02	−0.01	0.17	0.06

Depicts correlations between cortisol responses—i.e. AUCI and baseline-to-peak changes (%)—and self-stated stress measures—i.e. stress before and during task and primary appraisals. Coefficients are estimated by treatment and period: tournament during the first (T1) or second (T2) block and fixed payment during first (F1) or second (F2) block. ***$p<0.01$, **$p<0.05$, *$p<0.1$

self-stated stress during the treatment. These findings indicate that stress perceptions that are measured before and during (retrospectively) stressful situations are able to predict and explain a substantial part of acute cortisol responses to such situations.

With respect to primary appraisals—challenge and threat—we find coefficients that are close to zero. One's indication of appraising the treatment as challenging or threatening therefore provides little information about AUCI. It follows that partial correlations between AUCI and primary appraisals are close to zero. Accordingly, this measure is not able to pick up any signals that cause cortisol responses.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Table 7 Random-effects regression of AUCI on multiple self-stated stress measures

Dependent: AUCI	*baseline*	*model1*	*model2*	*model3*	*model4*
Stress before task (std)		18.23**			
Stress during task (std)			23.89***		
Challenge appraisal (std)				1.05	
Threat appraisal (std)					0.30
Treatment (1 if tournament)	63.22***	45.29*	38.76*	62.19***	62.96***
Period (1 if block 2)	43.42**	46.09**	51.79**	42.39*	43.41**
Treatment X Period	3.51	−4.24	−6.94	4.35	3.51
Constant	−80.61***	−70.39***	−69.08***	−79.86***	−80.47***
N	192	192	192	192	192
R^2 (overall)	0.17	0.19	0.20	0.17	0.17

Denotes coefficients of a random-effects regression model in which AUCI is regressed on self-stated stress measures and treatment and period indicators. Stress measures are standardized. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table A5 in Appendix A shows regression results for the baseline-to-peak increase of cortisol where results are similar.

3.3 Stress responses and self-selection

Having shown that a tournament scheme elicits greater stress responses than a fixed payment scheme, and that perceived stress and cortisol responses are in synch during treatments, we asked next whether workers self-select into payment schemes based on their stress responses.

3.3.1 Acute cortisol responses

In this section we investigate whether incentive choice is influenced by cortisol responses to the fixed payment and tournament. To investigate whether stress responses affect incentive choice, we use a linear probability model and regress compensation choice on stress responses for the entire sample. In addition to a simple linear regression, we include subjects' self-assessment, tournament performance, and risk attitude in the regression model. Moreover, all estimated models include session dummies. In line with expectations and previous literature (see Dohmen & Falk, 2011), we find that subjects who perform better in the first tournament, who have a higher self-assessment and who are more risk tolerant are more likely to select the tournament.[33] Due to the novelty effect, cortisol responses to incentives

[33] In a separate regression (results unreported), we jointly estimate the effects of these variables on incentive choice for the complete sample of 93 subjects of which 54 selected the tournament. Notice 3 observations are dropped due to multiple switching in the lottery choice task for risk aversion. In line with prediction, higher self-assessment positively affects entering the tournament (β=0.64, p=0.006),; moreover, subjects who solved more multiplications correctly when they were forced to work for tournament incentives earlier in the experiment are more likely to select the tournament (β=0.003, p=0.22),

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

are potentially not comparable for subjects in FT and TF. Consequently, we add an interaction effect in the regression for the orders FT and TF. Subsequently the estimated interaction effects are used to produce coefficients and standard errors of stress responses in both randomization orders.[34]

In Table 8 we assess the impact of the AUCI on tournament choice and assume that the novelty effect is independent of compensation choice. Table 8 denotes coefficient estimates of $AUCI_T$, $AUCI_F$, and $\Delta\ AUCI = AUCG_T - AUCG_F$ where indexes T and F indicate the tournament and fixed payment condition, respectively. AUCI measures are divided by 1000 for readability of the estimates. Each column represents a separate regression and depicts stress effects for subjects in randomization TF and FT. The relation between $AUCI_T$ and tournament choice, for both treatment orders, is positive and insignificant. The same holds for $AUCI_F$. The association between the difference of $AUCI_T$ and $AUCI_F$—i.e. Δ AUCI—, and tournament entry is also positive and insignificant. The coefficient estimates decrease in absolute size if we include the above mentioned variable set of controls and remain insignificant.[35] The decrease in absolute coefficient size suggests a correlation between these variables and cortisol responses. If we regress tournament entry on baseline-to-peak cortisol changes, then we find similar results: Cortisol responses do not affect compensation choice. As mentioned above, we assume the novelty effect does not influence subjects' preferences for the tournament or the fixed payment. Appendix C relaxes this assumption and estimates the relation between an isolated cortisol response—free of novelty effects—and tournament choice. We obtain similar results.

3.3.2 Subjective stress and incentive choice

In addition to cortisol responses, we assess whether subjective stress has an effect on the choice of payment scheme. Table 9 denotes regression coefficients of a linear probability model in which tournament entry is regressed on self-stated stated stress before and during both treatments, and primary appraisals before both treatments. Moreover, we have taken the difference between stress measures elicited during different incentive schemes—e.g. "Δ stress before" is the difference between self-stated stress before the tournament and the fixed payment. All regressors are standardized.

Table 9 shows little dependence between self-stated stress and compensation choice. Only perceived stress and appraised threat before the fixed payment show

Footnote 33 (continued)
and finally, subjects who are more willing to take risks are more likely to enter the tournament (β=0.55, p=0.15). Session effects are small and insignificant. Results are shown in Table A6 in Appendix A.

[34] In addition to the association between stress and selection, one may consider the relationship between stress and performance. Even though we do not focus on this relationship, we depict the correlations between stress and performance in Table A7 in Appendix A. The table shows that the correlation between stress and performance in the tournament is positive but small and statistically indistinguishable from zero.

[35] With respect to the variable *Risk attitude*, a total of 3 subjects switch between the safe payment and lottery more than once; 2 in incentive order FT and 1 in incentive order TF. In estimations where risk attitude is used as a predictor, these subjects are excluded from the sample.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Table 8 Linear probability models of tournament choice on AUCI

Comp. choice	*1 if tournament*		*1 if tournament*		*1 if tournament*	
$AUCI_T$ (TF)	0.51 (0.56)	0.24 (0.52)				
$AUCI_T$ (FT)	0.50 (0.86)	−0.12 (0.81)				
$AUCI_F$ (TF)			0.36 (1.78)	0.24(1.70)		
$AUCI_F$ (FT)			0.08 (0.62)	0.03 (0.58)		
Δ AUCI (TF)					0.49 (0.57)	0.24 (0.54)
Δ AUCI (FT)					0.14 (0.54)	−0.07 (0.51)
Controls	No	Yes	No	Yes	No	Yes
N	96	93	96	93	96	93

Denotes coefficients of a linear probability model in which tournament entry is regressed on cortisol response measures $AUCI_T$ and $AUCI_F$ and its difference $\Delta\ AUCI = AUCI_T - AUCI_F$. Estimates for subjects in incentive order TF and FT are shown by adding an interaction effect for incentive order (TF or FT) and depicting the resulting regression coefficients for both. All regressions include session effects. Control variables include *Risk attitude*, *Productivity*, and *Self-assessment*. Standard errors are depicted in parenthesis. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

a negative inclination towards selecting the tournament. As Table 9 contains 12 hypothesis tests, the probability of type 1 errors increases. If we implement Bonferroni correction and adjust the p-values accordingly, then we cannot reject any null hypothesis that the correlation is 0 at a 10% significance level. Consequently, we are hesitant to give any meaning to the aforementioned (marginally) significant correlations.

Akin to the results for cortisol responses, we find no evidence that self-stated stress responses to treatments impact treatment choice. Greater stress responses, either measured by cortisol or by self-stated perceptions, do not lead to different compensation choices.

4 Concluding discussion

We have documented that performance pay can be a stressor. In particular, we have shown that tournament incentives induce greater acute stress responses, measured both by cortisol responses and self-reported perceptions of stress, than a fixed reward. Notably, we observe a stress reaction even in our laboratory setting where the absolute, maximum stakes do not exceed 40 euros and exposure to incentives is no longer than 10 min. Arguably, this stress experience is confined in comparison to actual workplace settings. Our results also indicate that acute stress is induced both in settings in which the outcome of the tournament was revealed only to the participant and when it was publicly announced, indicating that social-evaluative threat is not the only stressor of tournament incentives. Additionally, we find evidence that subjective self-assessments of perceived stress are significantly correlated with the physiological stress response in terms of elevated cortisol levels, suggesting

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Table 9 Linear probability models of tournament entry on self-stated stress and appraisals

Comp. choice	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*	*1 if T*
stress before T	−0.02											
stress before F		−0.09*										
Δ stress before			0.05									
stress during T				0.01								
stress during F					−0.03							
Δ stress during						0.03						
challenge T							0.08					
challenge F								0.00				
Δ challenge									0.06			
threat T										−0.07		
threat F											−0.10*	
Δ threat												−0.01

Reports regression coefficients of a univariate linear probability model which regresses tournament entry on self-states stress and primary appraisals. All regressors are standardized. ***$p<0.01$, **$p<0.05$, *$p<0.1$

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

that individuals can be aware of the physiological consequences of performance pay. These findings have implications for the design of incentive schemes.

First, an important motivation for this study was to investigate whether tournament incentives can evoke an acute stress response, as evidence of a causal impact of incentives on acute stress is a prerequisite for a causal impact of performance pay on chronic stress. Chronic stress, i.e. exposure to stress over longer periods of time, is detrimental to physiological and psychological health, contributing to cardiovascular disease, metabolic syndrome, diabetes mellitus and mental disorders (see e.g. Cohen et al., 2007; McEwen, 2008). Our findings of an acute stress response suggest that performance pay potentially induces chronic stress and, eventually bad health. Clearly, our results can only be a first step towards assessing whether performance pay can be a pathway to adverse health conditions. It is conceivable, for example, that workers build resilience to working conditions when exposed to them over longer periods of time. Yet, our results point towards a promising research agenda to investigate the effects of incentive schemes at the workplace on stress and work-related health in the labor market. Another avenue for research would be to investigate whether our results also hold for women. Literature on stress responses has shown that gender might play a significant role in the reaction to stress. Where men are thought to react by the traditional theory of "fight or flight" (Cannon, 1932), more recent research suggests that women adhere to the so-called "tend and befriend" theory which postulates that women tend to join up with others for shared protection and comfort. In our experimental set up this could imply that women would decide to join tournaments significantly less, and restrain from competition. Previous experiments can also shed a light on to what would happen if we would have included women in the sample. Both Buser et al. (2017) and Zhong et al. (2018) find that in absence of an additional external stressor, which is closest to our setting, there are no gender differences in stress responses to competitive settings.

Second, the significant relationship between cortisol and self-reported responses provides insight into how anticipated stress can influence an individual's decision to enter stressful situations. If individuals experience and perceive stress when a specific stressor induces a cortisol response, they can incorporate the perceived signals about their hormonal response to that stressor in their decision-making process, even if that signal is imprecise. If individuals are aware of their hormonal stress response and gauge its costs adequately, they will make more informed and hence better decisions. For example, individuals might be induced to avoid the stressor, or at least persistent exposure to the stressor, if they estimate the costs of stress to be sufficiently large. Likewise, they could decide to expose themselves to the stressor if their expected rewards exceed the costs of stress.

A good stress assessment does not only improve an individual's decision-making, but also makes it feasible for researchers to ask individuals about their perceived stress, which can be an important basis for the design of policies. Our results show that up to 45 percent of the tournament effect on cortisol is explained by self-reported stress, which indicates that self-assessed stress is an adequate stand-in measure when cortisol measures are not available or too costly to obtain. An important caveat is that we could only assess the correlation between self-rated stress and

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

cortisol measures in one particular situation. Future research would have to show how generalizable our results are for other contexts.

Studying the relationship between cortisol and emotional stress responses has also attracted much scientific interest. Campbell and Ehlert (2012) discuss several studies that document such correspondence for studies that use the Trier Social Stress Test (TSST) as a stressor.[36] Approximately 25 percent of the reviewed studies show a statistically significant association between biological processes and subjective emotional responses. The authors suggest that dependencies might differ by subpopulations, timing of the self-reports, type of subjective stress measure and differences in adaptive mechanisms to stress responses. This study shows that in a homogeneous sample—e.g. male students—a relatively simple measure of emotional subjective stress—e.g. "To what extent do you feel stressed/calm/nervous" —is able to capture significant linear associations. Moreover, measuring self-reported stress at different points in time captures multiple dependencies: both anticipatory stress and perceived stress during the task are related to cortisol responses. In addition to the abovementioned features, the nature of the stressor may affect the extent to which the relationship between biological and subjective emotional stress manifest itself. For example, cortisol responses documented here are less volatile in comparison to responses to TSST. At the same time, this study's average cortisol response is small in comparison to responses to TSTT indicating a smaller signal of the stressor effect. Examining how the correspondence between biological and subjective emotional stress responses differs by stressors can shed further light on the measurability and existence of such correspondence.

Interestingly, we find that neither elevated cortisol levels nor self-perceived levels of stress affect the choice between tournament incentives and a fixed payment scheme, indicating that the costs or benefits of acute stress are not considered as factors in the sorting decisions. There are several potential explanations for this finding. First, perceived costs of stress associated with tournament entry might be low relative to its potential benefits. On the one hand, the tournament lasts only for a short period (i.e. 10 min) and is transient so that subjects might judge the absolute costs of stress to be low. On the other hand, potential benefits of playing the tournament may appear to be relatively large as participants can gain approximately €30 relative to the fixed payment if they win the tournament, but would only forgo about €10 if they lose the tournament. Hence, subjects may be willing to accept the costs of stress and are driven more by subjective winning probabilities to maximize earnings. Second, while individuals may shy away from the tournament because they perceive it as a threatening stressor, Gaab et al. (2005) propose that cortisol levels might also increase when appraising an event as challenging. Therefore, individuals who find it exciting to take on the challenge in the tournament might enter the tournament with elevated cortisol levels. As a result, greater cortisol responses may be

[36] "The Trier Social Stress Test […] has become one of the most popular methods for the experimental induction of acute psychosocial stress. Its highly standardized protocol comprises the delivery of a free speech for a job interview and an orally performed serial subtraction in front of an audience of supposed experts." (Ehlert and Campbell, 2012).

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

associated with both entering and rejecting the tournament. Finally, costs of stress may not be perceived correctly and hence not appropriately weighed in the sorting decision. With hindsight, some subjects might regret that they entered the tournament because the stress it induces may impact them negatively later during the day. At the work place, stress is likely to be experienced repeatedly and over longer periods so that individuals become acquainted with the effects of stress. But even then, individuals might not perceive the long-term consequences of stress accurately, particularly because potential negative effects of stress may initially not be felt. Moreover, sufficiently impatient individuals might ignore the long-term costs of stress, which accrue in the future in the form of deteriorated health. Assessing the role of these factors for sorting decisions in the work place is worth further investigation.

For the design of optimal policy it would be important to better understand what considerations about stress and its costs affect the sorting decision.[37] For example, if individuals deliberately expose themselves to a stressor but do not take detrimental effects of stress into account, there is potentially scope for policy intervention. Of course, acute stress may affect sorting differently than chronic stress as individuals prepare themselves for a short exertion of effort in a one-period game. In our setup, we cannot study whether workers become aware of chronic stress when they are exposed to the stressor for an extended period, and whether they would respond to chronic stress experience by avoiding tournaments. Nevertheless, answering this question is a challenge for future research.

Finally, from a methodological perspective, our finding of a novelty effect—i.e. participants tend to be stressed at the beginning of the experiment even in absence of any exogenously induced stressor—has some important implications for the interpretation of experimental data and the design of experiments. Greater cortisol levels, or HPA activation in general, are associated with important cognitive and affective processes. Depressive symptomology has been associated with increased cortisol levels (Heim & Nemeroff, 1999). Moreover, heightened HPA activity can have effects on memory (e.g. Buchanan & Lovallo, 2001). Consequently, information that is extracted from experiments may be affected by initial stress responses. Another important observation is that novelty efxfects differ by individuals. Consequently, choices during experiments or surveys may be driven by differences in novelty effects. These considerations have implications for the design of experiments. In experiments like ours, in which cortisol responses are measured, but also in experiments across the board, it is advisable to start with a "cool-down"-phase before the main experiment starts, in order to rule out that experimental results are confounded by participants' stress reactions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-023-09801-1.

[37] A recent study by Nagler et al. (2022) sheds light on the considerations about stress and occupational sorting, and studies whether individuals require compensation for greater work pressure. The study shows that "work pressure comes with a sizeable earnings premium even within narrowly defined occupations." Further research is needed to analyze whether individuals then require a stress premium for their acceptance of pay-for-performance contracts.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Acknowlegements The authors gratefully acknowledge financial support from NWO Vidi grant 452–10-006. Dohmen and Rohde gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A05). Dohmen also acknowledges funding by DFG under Germany´s Excellence Strategy—EXC 2126/1– 390838866. The replication material for the study is available at https://doi.org/10.34894/9KL75I.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adam, E. K., Hawkley, L. C., Kudielka, B. M., & Cacioppo, J. T. (2006). Day-to-day dynamics of experience–cortisol associations in a population-based sample of older adults. *Proceedings of the National Academy of Sciences, 103*, 17058–17063.

Allan, J. L., Andelic, N., Bender, K. A., Powell, D., Stoffel, S., & Theodossiou, I. (2021). Employment contracts and stress: Experimental evidence. *Journal of Economic Behavior & Organization, 187*, 360–373.

Bozovic, D., Racic, M., & Ivkovic, N. (2013). Salivary cortisol levels as a biological marker of stress reaction. *Med Arch, 67*(5), 374–377.

Buchanan, T. W., & Lovallo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology, 26*(3), 307–317.

Buckert, M., Schwieren, C., Kudielka, B. M., & Fiebach, C. J. (2017). How stressful are economic competitions in the lab? An investigation with physiological measures. *Journal of Economic Psychology, 62*, 231–245.

Buser, T., Dreber, A., & Mollerstrom, J. (2017). The impact of stress on tournament entry. *Experimental Economics, 20*, 1–25.

Cahlíková, J., Cingl, L., & Levely, I. (2020). How stress affects performance and competitiveness across gender. *Management Science, 66*(8), 3295–3310.

Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: Does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology, 37*, 1111–1134.

Cannon, W. B. (1932). *The wisdom of the body*. Norton.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods., 4*, 62–83.

Cohen, S., Janicki-Deverts, D., & Miller, G. E. (2007). Psychological stress and disease. *JAMA, 298*, 1685–1687.

Cohen, S., Schwartz, J. E., Epel, E., Kirschbaum, C., Sidney, S., & Seeman, T. (2006). Socioeconomic status, race, and diurnal cortisol decline in the coronary artery risk development in young adults (CARDIA) study. *Psychosomatic Medicine, 68*, 41–50.

Costa, P., & McCrae, R. (1992). Four ways five factors are basic. *Personality Individual Differences., 13*, 653–665.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*, 355.

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

EU-OSHA (2013). Pan-European opinion poll on occupational safety and health. *Publications Office of the EU*.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Eurofound and EU-OSHA (2014). Psychosocial risks in Europe: Prevalence and strategies for prevention. *Publications Office of the European Union*.

Eysenck, M. W. (1992). *Anxiety: The cognitive perspective*. Psychology Press.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*, 171–178.

Fletcher, G. J., Danilovics, P., Fernandez, G., Peterson, D., & Reeder, G. D. (1986). Attributional complexity: An individual differences measure. *Journal of Personality and Social Psychology, 51*, 875.

Folkman, S. (1984). Personal control and stress and coping processes: A theoretical analysis. *Journal of Personality and Social Psychology, 46*, 839.

Gaab, J., Rohleder, N., Nater, U. M., & Ehlert, U. (2005). Psychological determinants of the cortisol stress response: The role of anticipatory cognitive appraisal. *Psychoneuroendocrinology, 30*, 599–610.

Gruenewald, T. L., Kemeny, M. E., Aziz, N., & Fahey, J. L. (2004). Acute threat to the social self: Shame, social self-esteem, and cortisol activity. *Psychosomatic Medicine, 66*(6), 915–924.

Heim, C., & Nemeroff, C. B. (1999). The impact of early adverse experiences on brain systems involved in the pathophysiology of anxiety and affective disorders. *Biological Psychiatry, 46*(11), 1509–1522.

Houston, J., Harris, P., McIntire, S., & Francis, D. (2002). Revising the competitiveness index using factor analysis. *Psychological Reports, 90*, 31–34.

Izard, C. E. (1993). *The Differential Emotions Scale: DES IV-A: a Method of Measuring the Meaning of Subjective Experience of Discrete Emotions*. University of Delaware.

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology, 19*(4), 313–333.

Kirschbaum, C., Prussner, J. C., Stone, A. A., Federenko, I., Gaab, J., Lintz, D., Schommer, N., & Hellhammer, D. H. (1995). Persistent high cortisol responses to repeated psychological stress in a subpopulation of healthy men. *Psychosomatic medicine, 57*, 468–474.

Kirschbaum, C., Strasburger, C. J., & Langkrär, J. (1993). Attenuated cortisol response to psychological stress but not to CRH or ergometry in young habitual smokers. *Pharmacology Biochemistry and Behavior, 44*, 527–531.

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty impedes cognitive function. *Science, 341*, 976–980.

McEwen, B. S. (2008). Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators. *European Journal of Pharmacology, 583*, 174–185.

Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: Evaluation of salivary cortisol pulse detection in panel designs. *Psychosomatic Medicine, 75*(9), 832–840.

Molitch, M. E. (1995). Neuroendocrinology. In P. Felig, J. D. Baxter, & L. A. Frohman (Eds.), *Endocrinology and metabolism* (3rd ed., pp. 221–275). McGraw-Hill Inc.

Nagler, M., Rincke, J., & Winkler, E. (2022). High-pressure, high-paying Jobs?. *Available at SSRN 4291736*.

Nolte, H., Weischer, C., Wilkesmann, U., Maetzel, J., & Tegethoff, H. G. (1997). Kontrolleinstellungen zum Leben und zur Zukunft–Auswertung eines neuen sozialpsychologischen Itemblocks im Soziooekonomischen Panel. *Diskussionspapier der Fakultät für Sozialwissenschaft der Ruhr-Universität Bochum*.

Pruessner, J. C., Kirschbaum, C., Meinlschmid, G., & Hellhammer, D. H. (2003). Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology, 28*, 916–931.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*, 1.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the life orientation test. *Journal of Personality and Social Psychology, 67*, 1063.

Smyth, J. M., Ockenfels, M. C., Gorin, A. A., Catley, D., Porter, L. S., Kirschbaum, C., & Stone, A. A. (1997). Individual differences in the diurnal cycle of cortisol. *Psychoneuroendocrinology, 22*, 89–105.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press

Spybrook, J., Bloom, H., Congdon, R., Hill, C., Martinez, A., Raudenbush, S., & To, A. (2011). Optimal design plus empirical evidence: Documentation for the "Optimal Design" software. *William T. Grant Foundation. Retrieved on November*, *5*, 2012.

Stone, A. A., Schwartz, J. E., Smyth, J., Kirschbaum, C., Cohen, S., Hellhammer, D., & Grossman, S. (2001). Individual differences in the diurnal cycle of salivary free cortisol: A replication of flattened cycles for some individuals. *Psychoneuroendocrinology, 26*(3), 295–306.

Weitzman, E. D., Fukushima, D., Nogeire, C., Roffwarg, H., Gallagher, T. F., & Hellman, L. (1971). Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. *The Journal of Clinical Endocrinology & Metabolism, 33*, 14–22.

Zhong, S., Shalev, I., Koh, D., Ebstein, R. P., & Chew, S. H. (2018). Competitiveness and stress. *International Economic Review, 59*(3), 1263–1281.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.1007/s10683-023-09801-1 Published online by Cambridge University Press