Overexertion of Effort under Working Time Autonomy and Feedback Provision

Citation for published version (APA):

Dohmen, T., & Shvartsman, E. (2023). Overexertion of Effort under Working Time Autonomy and Feedback Provision. *Journal of Economic Behavior & Organization*, *212*, 1255-1266. https://doi.org/10.1016/j.jebo.2023.05.043

Document status and date:
Published: 01/08/2023

DOI:
10.1016/j.jebo.2023.05.043

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Journal of Economic Behavior and Organization 212 (2023) 1255–1266

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Overexertion of Effort under Working Time Autonomy and Feedback Provision[☆]

Thomas Dohmen[a], Elena Shvartsman[b,*]

[a] University of Bonn and IZA Germany; Maastricht University, The Netherlands
[b] WHU - Otto Beisheim School of Management and IZA Germany

ARTICLE INFO

Article history:
Received 12 October 2021
Revised 12 May 2023
Accepted 28 May 2023

JEL classification:
C91
D90
I10
J81

Keywords:
working time autonomy
performance uncertainty
feedback provision
incentives
effort
subjective stress

ABSTRACT

Working time autonomy is often accompanied by output-based incentives to counterbalance the loss of monitoring that comes with granting autonomy. However, in such settings, overprovision of effort could arise if workers are uncertain whether their performance suffices to secure the output-based rewards. Performance feedback can reduce or eliminate such uncertainty. We demonstrate in a laboratory experiment that a precautionary effort motive can lead to overprovision of costly effort in work environments with working time autonomy in the absence of feedback. A key feature of our design is that it allows for a clean measurement of effort overprovision by keeping performance per unit of time fixed, which we achieve by calibrating subjects' productivity on a real effort task ex ante. This novel design can serve as a workhorse for various experiments as it allows for exogenous variation of performance certainty (i.e., by providing feedback), working time autonomy, productivity, effort costs, and the general incentive structure. We find that subjects provide significantly more costly effort beyond a level necessary to meet their performance targets in the presence of uncertainty, i.e., the absence of feedback, which suggests that feedback shields workers from overprovision of costly effort.

1. Introduction

Worker autonomy becomes increasingly important in the era of digitalization, where information technologies allow for a higher degree of decentralization in many jobs (Gibbs, 2017). A pivotal aspect of autonomy is working time autonomy, i.e., the discretion with respect to when or (and) how much to work (see, e.g., Beckmann et al., 2017), which is usually

☆ We would like to thank Daniela Puzzello, two anonymous reviewers, Simone Quercia, and conference and seminar participants at the 24th Colloquium on Personnel Economics (COPE), the Lüneburg Workshop in Microeconomics, the WHU – Otto Beisheim School of Management, and the University of Trier for their useful comments and discussions. Elena Shvartsman acknowledges financial support from the Swiss National Science Foundation under grant No. IZSEZ0 177659/1. Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A05) and under Germany's Excellence Strategy EXC 2126/1-390838866 is gratefully acknowledged. We would like to thank Maximilian Blesch, Jana Hofmeier, Thomas Neuber, and Rafael Suchy for excellent research support. All remaining errors are our own. This study is based on four treatments of an experiment described in the AEA RCT Registry (#0003532).

* Corresponding author.
E-mail addresses: tdohmen@uni-bonn.de (T. Dohmen), elena.shvartsman@whu.edu (E. Shvartsman).

https://doi.org/10.1016/j.jebo.2023.05.043

embedded in a bundle of other discretionary policies such as working from home (e.g., Kelly and Moen, 2007).[1] Autonomy is probable to be granted in uncertain environments in which, at the same time, output-based performance is likely to occur; for principals would tend to delegate responsibility to workers and grant them discretion about how and when to work if the environment is uncertain (Prendergast, 2002). At the same time, actual working time and effort are not easily monitored in such settings, so that 'output-based incentive pay is more likely to be observed' (Prendergast, 2002, p. 1072) in settings where workers have working time autonomy. An example of output-based incentives is a bonus payment for reaching a performance target. In work settings, where workers are uncertain about whether they are reaching a performance target and have the discretion to provide more effort, workers might work too much, for instance, due to precautionary effort provision (Eeckhoudt et al., 2012) or because they are loss averse (Corgnet and Hernán-González, 2019; Sloof and Van Praag, 2010).[2]

In this study, we document in a stylized laboratory setting that working time autonomy characterized by performance uncertainty induces overprovision of personally costly effort, i.e., effort beyond the level necessary to achieve a specific performance target. We operationalize overprovision of effort by additional costly working time (Bell and Freeman, 2001). We test whether reducing performance uncertainty by the provision of feedback curbs the overexertion of effort. In addition, we investigate whether and under what conditions working time autonomy induces subjective stress.

Our study makes two contributions to the literature. First, we complement studies on working time autonomy and effort provision. These studies rely on field data to show that individuals provide more effort in terms of working hours under working time autonomy (Avgoustaki and Bessa, 2019; Beckmann et al., 2017), in particular when working time autonomy is combined with performance-related pay (Lott, 2014). However, the role of performance uncertainty under working time autonomy and the potentially mitigating effects of feedback provision have not yet received attention. In this study, we, therefore, contrast performance uncertainty with certainty by implementing performance feedback in our experimental conditions. Moreover, establishing the causal nature of the relationship between working time autonomy under uncertainty and (suboptimal) provision of overtime work from observational data is complicated by endogeneity problems due to omitted variables, worker selection, or measurement error. Hence, with this study, we expand our understanding of effort provision in an environment characterized by working time autonomy and performance uncertainty.

Second, our study adds to the literature on feedback, in which feedback is usually used to inform individuals about their relative ranking within a peer group. This literature studies how relative performance feedback affects performance and productivity when performance is related to pay (Azmat and Iriberri, 2016; Eriksson et al., 2009) or when individuals might receive utility from their position in a ranking (Azmat et al., 2019; Blanes i Vidal and Nossol, 2011). Unlike these studies, we provide feedback about individuals' absolute performance, not their relative performance. Moreover, in our study, feedback does not affect output through changes in productivity, i.e., output per unit of time, because we control productivity by design. In our setting, feedback may influence effort decisions by decreasing (or eradicating) uncertainty about one's productivity. So, while the above-cited studies have shown that relative feedback can lead to higher output and effort provision, our study highlights how feedback can shield workers from overexertion of effort.

We designed a laboratory experiment to assess whether working time autonomy coupled with output-based incentives can lead to overprovision of effort in the presence of performance uncertainty and whether the provision of feedback, which reduces performance uncertainty, curbs such overexertion of effort. In a 2×2 design, we varied whether working time autonomy was granted and whether feedback on performance was provided. The four conditions resulting from our 2×2 design are *AutonomyFeedback, AutonomyNoFeedback, NoAutonomyFeedback*, and *NoAutonomyNoFeedback*. In the *Feedback* conditions, we informed participants constantly about their performance and hence their current distance to the performance target. In the *NoFeedback* conditions, participants received information about their productivity only at the end of the experiment and hence remained uncertain about their performance at the end of the work task. In all conditions, participants worked on a real effort task, in which they had to estimate the number of blue dots on a black screen. During the regular working time of approximately 9 minutes, participants could work on 45 units of this estimation task. In order to achieve the performance target which made workers eligible for a bonus payment, they had to solve 34 estimation problems correctly.

In the *Autonomy* conditions, participants were free to prolong working time at a cost or stop working at any time. Also, participants knew at the start of the work phase that they would be able to step-wise purchase batches of additional working time, thereby forgoing a part of their bonus. We define overtime as any additional batch that subjects decided to work on above the regular working time. The real effort task allows us to measure overtime provision precisely and calibrate expected productivity, i.e., number of correct answers per minute, ex ante. We thereby control participants' productivity. We calibrated the task such that approximately 90 percent of all participants reached the performance target irrespective of the condition. This productivity made overprovision of effort, which equaled extra work in our context, inefficient for almost all participants. Hence, our experimental manipulation allows us to observe whether certain working conditions provoke overprovision of costly effort. We also measured self-assessed stress during the experiment to elicit whether and under what conditions working time autonomy deteriorated subjective well-being.

[1] In many European countries, a substantial fraction of the workforce has a high degree of working time autonomy. In the sample of 35 countries covered by the European Working Conditions Survey 2015, for example, the median percentage of individuals per country that answered 'Working hours entirely self-determined' when asked 'How are your working time arrangements set?' is 16 percent (see, EWCS, 2016).

[2] This dilemma of uncertainty and overwork spurred by the thought that one could always 'do more' received recent attention in a Twitter discussion (Fiesler, 2021).

Our design has several attractive features. First, we can rule out self-selection into particular working conditions, allowing us to identify a causal effect of working conditions on effort provision. Second, we can implement experimental conditions that contrast effort provision under certainty about performance and performance uncertainty. We manipulate performance uncertainty by the provision of feedback. Third, our framework allows us to elicit different outcomes associated with varying working conditions, such as perceived stress. Lastly, we employ a real effort task that allows us to calibrate subjects' performance. These design features are also valuable when addressing different questions regarding the relationship between autonomy, effort provision, and work-related outcomes under constant productivity. Moreover, our design permits manipulations beyond the features we vary in this application, such as effort costs or task productivity. Hence, our experimental setup can be a workhorse for studying various research questions pertaining to working time autonomy.

We find that participants in our experiment provide more costly effort in the condition *AutonomyNoFeedback* than in the condition *AutonomyFeedback*. This finding indicates that absent feedback, i.e., under uncertainty about one's productivity, the combination of permitting subjects to work longer (working time autonomy) and offering output-based incentives causes unnecessary overtime work from an ex post perspective, reducing subjects' financial welfare. However, we do not find that working time autonomy, irrespective of whether feedback is provided, translates into increased self-assessed stress in our setting, where total working time is arguably short. If anything, we find that working time autonomy mildly reduces perceived stress.

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

2. Experimental Design

We designed an experiment in which subjects work on a real effort task and receive a bonus paid if they reach an absolute performance target. In a 2×2 design, we vary whether subjects receive working time autonomy and whether they receive feedback about their performance. The first implies that subjects can decide to deviate from the regular working time of 9 minutes by stopping earlier or buying additional batches of the real effort task. The latter allowed us to vary whether subjects are certain or uncertain about their productivity.

2.1. Experimental Task

The real effort task that subjects had to work on is adopted from Fliessbach et al. (2007). Fig. 1 displays the task. In this task, subjects see blue dots randomly distributed on a black screen for 1.5 seconds. On the next screen, they see a number for seven seconds and have to decide whether they have seen more or fewer dots than the suggested number. We inform subjects about the time left for their decision via a countdown in the upper right corner of the decision screen. During the experiment, answers would be automatically recorded as wrong if no decision was submitted within the seven seconds time frame. In the *Feedback* conditions, we also informed subjects about the cumulative number of correctly solved pictures in the upper left corner of each decision screen. Subjects saw the subsequent screen for 2.5 seconds. This screen either provided feedback on the previous unit (whether it was correct or wrong) or contained the information that the work is about to proceed in the *Feedback* conditions and *NoFeedback* conditions, respectively. Before the work phase, all subjects worked on five units of the task without receiving performance feedback.

This estimation task has various advantageous features for our study purposes. First, the task is divisible into smaller blocks of fixed duration, which permits us both to offer subjects extra units and to allow them to stop working at any time. Crucially, we kept the duration of each working unit, i.e., one estimation problem, and each batch, which comprises five estimation problems, constant by displaying a new task on the screen every eleven seconds. Since workers cannot influence the pace at which a new task starts, they can only work on more tasks by increasing working time. The choice of working time, i.e., the number of batches that participants choose to work on, therefore corresponds to participants' effort choice. Hence, we can infer from our subjects' working time choices how much effort they provided.[3] Moreover, we can exogenously vary the marginal cost of working by varying the amount workers have to pay in order to work on an additional unit of the real effort task. In our experiment, we set this cost to zero for the first 45 units of the task (i.e., the regular working time), and then implement a positive but constant marginal cost for additional estimation tasks during overtime (see below). In addition to this exogenous component of marginal effort costs, the task is tedious and therefore induces disutility.

Second, there are only minor productivity differences across subjects in this task because capacity limits to human perception largely determine task productivity. Therefore, differences in ability and effort across subjects are controlled for by design. In fact, a comparison of the coefficient of variation of performance across different widely used real effort tasks reveals that performance heterogeneity in our real effort task is low. Table A.1 in Appendix A shows that the coefficient of variation for the task used in our experiment is 0.106, while the respective coefficients of variation for all other tasks, which range from 0.173 for the encryption task (Erkal et al., 2011) to 0.555 for a multiplication task (Dohmen and Falk, 2011), are larger, with the exception of Kuhn and Villeval (2015).[4]

[3] Our design, therefore, entails that effort is defined as overtime work, but not as work intensity (see, Avgoustaki and Frankort, 2019, for a definition).

[4] Furthermore, a regression of task productivity (number of correct answers during the regular working time) on cognitive skills assessed via the final mathematics grade in school and performance in the Raven matrix test during the experiment did not reveal any statistically significant associations.

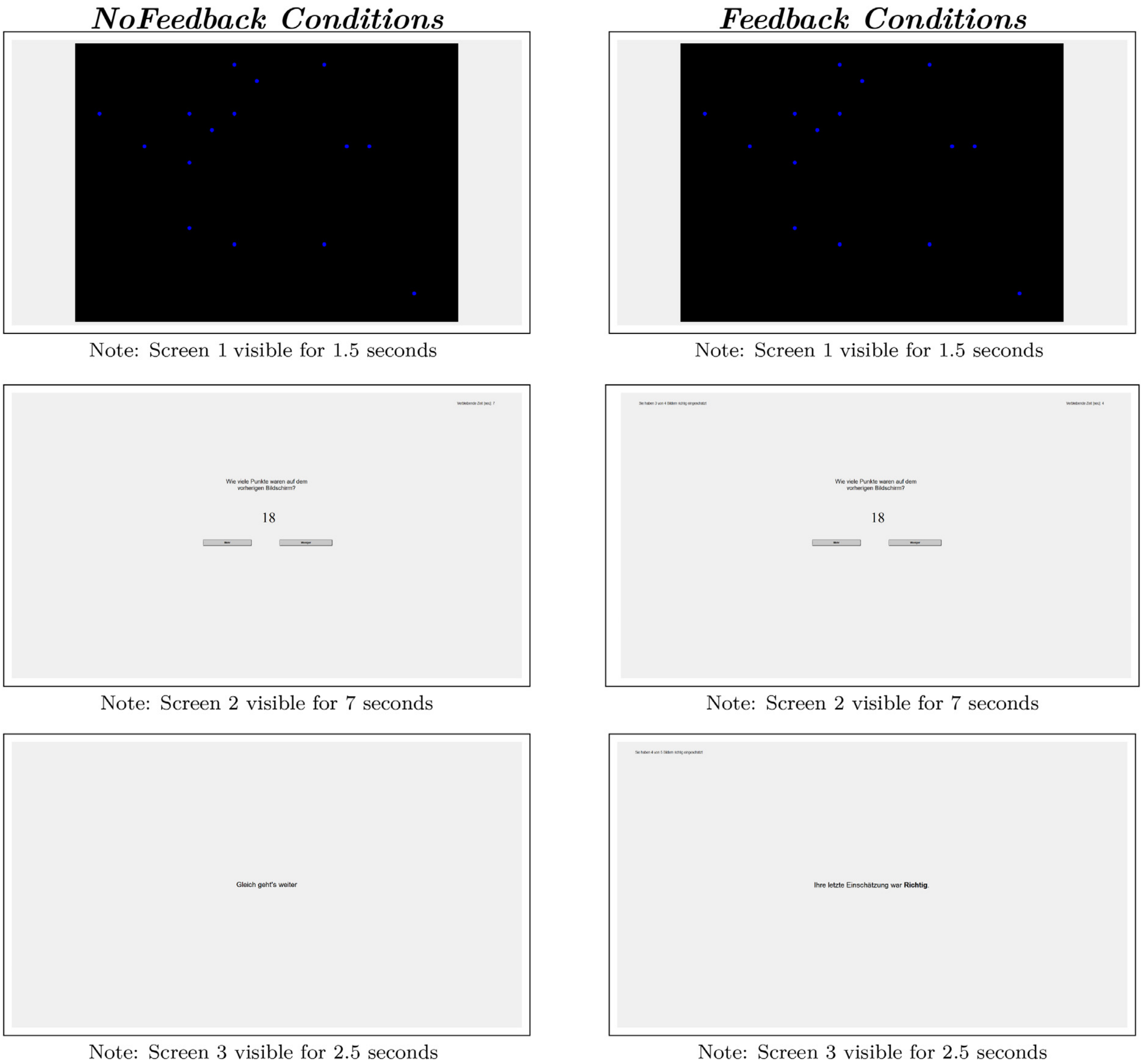

Fig. 1. Experimental Task
Note: This image displays screenshots of the original experiment in German, please see the Online Appendix for a translation into English.

Third, productivity can be calibrated ex ante by selecting the number of dots visible on the first screen and the suggested number of dots on the following screen. We generated pictures with four to 55 randomly distributed dots for this task. We chose four to 55 dots per picture in line with Fliessbach et al. (2007) and followed them by deviating the suggested number of dots by approximately 20 percent. The number of dots visible on the screen and the deviation of the suggested number of dots from the actual number allowed us to vary the difficulty level per estimation problem. We then conducted a calibration study to estimate the average productivity for each picture.[5] We used the data collected in the calibration study, i.e., the rate of correct estimates per picture, to construct work sequences that include 45 estimation tasks such that approximately 90 percent of our subjects are expected to reach the performance target of 34 correct answers during the regular working time. It turns out that 90.3 percent of our subjects achieve the performance target during the regular working time.

Overall, these features of the task enabled us to ensure ex ante that a certain percentage of our participants achieve the performance target during the scheduled working time and that this percentage is essentially the same across conditions.

[5] This calibration study was conducted at the Laboratory for Experimental Economics at the University of Bonn (BonnEconLab) in October 2018 with 44 participants, whom we recruited from the BonnEconLab subject pool via hroot (Bock et al., 2014). Most participants were students (90.9% students), aged on average 24.95 years, and 50 percent of the participants were female. Participants worked for approximately 50 minutes at the computer and earned on average EUR 13.87. They received a piece rate of five experimental points (exchange rate of 100=EUR 1) per correct estimate. In order to reduce potential fatigue effects, we gave participants a relaxation break after half of the working time.

Most importantly, however, we thereby ensure that for most subjects, provision of extra working time is inefficient, which allows us to quantify overprovision of effort.

2.2. Treatment Conditions

Subjects were randomly assigned to four conditions in a 2×2 design that varied assignment to *Feedback* or *NoFeedback* treatment and assignment to working under a working time autonomy regime, *Autonomy*, or a fixed working time schedule, *NoAutonomy*. This resulted in four conditions, *AutonomyFeedback, AutonomyNoFeedback, NoAutonomyFeedback*, and *NoAutonomyNoFeedback*. In all four conditions, we asked subjects to work on 45 units of the real effort task corresponding to roughly 9 minutes, i.e., the regular working time.

In the *Autonomy* conditions, we grant workers autonomy over their working time, which is the possibility to choose their working time duration. Specifically, we offered subjects the opportunity to prolong their working time by buying additional batches of five units of the real effort task at the cost of 40 points per batch. Also, we explicitly informed subjects that they could stop working anytime. In order to create opportunity costs of working during the work phase, subjects could read magazines when not working.[6]

In the *Feedback* conditions, we provide constant feedback on productivity throughout the working phase as subjects are always informed about the number of correctly solved estimation tasks. In contrast, we did not inform subjects about their productivity until the end of the experiment in the *NoFeedback* conditions. In all conditions, subjects knew they needed to submit at least 34 correct estimates – disregarding the actual duration of their work – to reach a bonus of 360 experimental points.

The main objective of this study is to investigate whether feedback provision in conditions where workers have working time autonomy shields them from providing costly additional working time. We, therefore, predominantly compare and analyze the two *Autonomy* conditions as these allow subjects to choose their working time. Particularly, our design allows us to causally identify the effect of uncertainty, i.e., the absence of feedback, on the provision of overwork under working time autonomy.

We also implemented two further stages to study sorting decisions into working time regimes. In stage two, subjects worked under the opposite working time regime compared to stage one, i.e., workers under *Autonomy* now faced fixed working time and vice versa. In the third stage, subjects could choose to work under *Autonomy* or under a fixed working time regime. In this paper, we focus on data from the first stage of the experiment as the first stage allows for a clean comparison of effort choice.

2.3. Additional Measures

Apart from recording their working time choices, we assessed participants' appraisal of a situation's stressfulness via the primary appraisal index, which we calculated from eight items of the standardized PASA questionnaire (Gaab et al., 2005).[7] Participants answered the respective questions before task onset to measure subjective anticipatory stress. We also surveyed subjective stress perceptions and other feelings before and after the task (Dohmen and Falk, 2011; Dohmen et al., 2023). The respective statement referring to stress before task onset is 'Right now I feel stressed' with an affirmation scale ranging from 'not at all' (1) to 'completely' (5). We also asked subjects how calm, tense, satisfied, and tired they felt. The respective statement for post task stress evaluation is 'How stressed did you feel?', where again, the level of affirmation ranged from 'not at all' (1) to 'completely' (5). Here, we also asked subjects how much effort they exerted and about their feelings of exhaustion, calmness, tension, satisfaction, and tiredness. In addition, we surveyed demographics, measured cognitive skills (assessed via a Raven matrix test and the final mathematics grade in school), and elicited subjects' risk preferences using self-assessed willingness to take risk (general and domain-specific, see, Dohmen et al., 2011) and an incentivized lottery choice (multiple price list format). Finally, for exploratory reasons, we elicited further personality traits, attitudes, and preferences after the dependent variable (DV), which we did not use in this study; these are listed in the Online Appendix. We describe the chronological sequence of the entire study in Appendix A.

2.4. Procedures

We conducted our experiment at the Laboratory for Experimental Economics at the University of Bonn (BonnEconLab) in November 2018.[8] We recruited participants from the BonnEconLab subject pool via hroot (Bock et al., 2014); 175 subjects

[6] A set of magazines was distributed to each participant before the experiment. Instead of the typical outside option of browsing the web (Corgnet et al., 2015; Goerg et al., 2019), we allowed subjects to read magazines of neutral content to avoid that our outcome variables with respect to subjective stress perceptions were compromised by potential excitement of uncontrolled web browsing.

[7] In specific, we first calculated the primary appraisals index' sub-constructs, challenge and threat, from four items each and averaged these to form the primary appraisals index.

[8] All experimental procedures were carried out in accordance with the rules for conducting experiments at BonnEconLab. This study is based on four treatment arms of an experiment that is described in the AEA RCT Registry (#0003532). We pre-registered six treatment conditions, the four treatment conditions described in Section 2.2, which all entailed output-based incentives, and two supporting conditions with fixed working time and fixed compen-

participated in eight sessions.[9] Most of the subjects were students (94.9%). Their average age was 24.5 years, and 63.4 percent of participants were female. Subjects spent around two hours at the laboratory and earned EUR 26.33 on average. We randomized subjects to treatment conditions at the session level. Subjects were paid privately at the end of each session.

The experiment was programmed in zTree (Fischbacher, 2007), and all instructions were in German (see the Online Appendix for the instructions and screenshots of the experiment). We informed subjects that they would receive their final payout in Euro but that we would refer to points at an exchange rate of 100 points per Euro throughout the experiment.

We check whether subjects are balanced across conditions by regressing key variables such as age, gender, and cognitive skills on indicator variables for the conditions. We do not find any statistically significant differences, indicating that subjects were similar across conditions with respect to their observable characteristics. We report the results of the associated joint F-tests as well as means and standard deviations of the inspected key variables per condition in Table A.2 in Appendix A.

2.5. Hypothesis

Our hypothesis is that working time autonomy, which allows workers to provide extra working time, leads to overprovision of effort if remuneration depends on a performance target and that the effect is moderated by the degree of uncertainty about reaching this target. Overprovision of effort could result because of precautionary effort in the presence of uncertainty about securing future bonus payments (Eeckhoudt et al., 2012). Prudent workers dislike downside risk. As a result, they strive to reduce the risk that their actual performance is below the target. In our setup, they can do so by working longer. Under uncertainty, workers choose working time before actual performance is revealed, which induces precautionary effort. This motive for precautionary effort provision disappears if workers know their actual performance with certainty at the time when workers can decide to stop working or to prolong working time.

Another motive for overprovision of effort in uncertain environments derives from loss aversion. If workers are loss-averse with respect to a performance target, they are willing to work longer to avoid the loss they would experience by not achieving the performance target and hence not receiving the bonus payment. Loss aversion as a driver of higher effort provision in more uncertain environments is discussed in Sloof and Van Praag (2010) and Corgnet and Hernán-González (2019).[10]

Feedback provision can curb these motives as it reduces uncertainty. Once subjects learn that they have reached the performance target, these motives should even become irrelevant. We, therefore, expect that subjects will more likely use the option of buying additional working time in the absence of feedback:

Hypothesis 1: *Subjects in the AutonomyNoFeedback condition will buy more overtime than subjects in the AutonomyFeedback condition.*

3. Results

3.1. Performance Uncertainty and Working Time Choices

We begin this section with our main results and analyze whether subjects extend their working time at a personal cost when being uncertain about their productivity. Recall that we designed the task such that 90 percent of subjects should have reached the performance target within regular working time so that we expect approximately nine of the 88 participants in the *Autonomy* conditions to fall short of the performance target after regular working time and to buy additional units of work. We find that four subjects in the *AutonomyNoFeedback* and five in the *AutonomyFeedback* condition did not meet the performance target after regular working time. However, many more decided to buy additional working time.[11] Twenty five subjects in *AutonomyNoFeedback* extended their working time, which are significantly more subjects than the twelve individuals who bought extra time in *AutonomyFeedback* (Wilcoxon-Mann-Whiney test, p-value = 0.024).[12] As a result, the

sation. We did not implement the two supporting conditions. In order to study the effects of working time autonomy under uncertainty and incentives on effort provision, the conditions with fixed working time and a fixed payout are irrelevant. For this reason, we focus on the four main conditions which implement output-based incentives.

[9] Note that we preregistered to recruit about 70 subjects in each treatment arm. This sample size was pre-registered in the absence of information about the minimum detectable effect size for main outcomes. Eventually, we decided to stop the data collection after having run two sessions for each treatment condition, i.e., after having collected data for 41 to 47 subjects per condition. Our ex-post power calculations, which are based on observed effect sizes and should therefore be regarded with caution, suggest that our sample is sufficiently powered (see, footnote [18]).

[10] Further potential explanations include, the misperception of effort costs due to bounded rationality (Simon, 1955; 1972). Also, a gift-exchange type of argument (Akerlof, 1982) is conceivable, i.e., agents could perceive the prospect of a bonus as something they wish to reciprocate by providing more effort. This explanation is closely linked to the idea that workers not only provide effort to obtain a reward for reaching a performance target but also to signal their aptitude to employers, be it in expectation of future job opportunities or because of reciprocity. Also, working time autonomy may augment intrinsic motivation (Beckmann et al., 2017) and therefore induce additional effort.

[11] While it would have also been interesting to investigate whether and under which conditions subjects use *Autonomy* to stop working before the end of the regular working time, there is little scope for such an analysis as only two subjects, both assigned to the *AutonomyFeedback* condition, stopped working prematurely.

[12] Only one out of the twelve participants who bought additional working time in treatment condition *AutonomyFeedback* had not achieved the performance target during regular working time. Answers to the final open field questionnaire of the other eleven participants suggest that they mostly either forgot the performance target or had not fully understood the incentive structure.

Table 1
Working Time Purchase Decisions.

Dependent variable	Purchase of Additional Working Time (Marginal Effects)		
	(1) (Probit)	(2) (Probit)	(3) (Probit)
AutonomyNoFeedback	0.239**	0.241**	0.276**
	(0.102)	(0.104)	(0.106)
Age		-0.005	-0.003
		(0.011)	(0.011)
Female		−0.199*	−0.188*
		(0.110)	(0.112)
Risk		0.021	0.021
		(0.023)	(0.024)
Productivity Test Round			0.016
			(0.364)
Raven Matrix Test			0.027
			(0.031)
Productivity Batch 1 and 2			-0.073
			(0.044)
Observations	88	88	88
Pseudo R^2	0.044	0.083	0.112

Notes: Marginal effects reported in all columns. The according estimated coefficients can be found in Table A.3 in Appendix A. Condition *AutonomyFeedback* serves as reference category. The values in parentheses represent standard errors. */**/*** denotes statistical significance at the 10/5/1% level. All specifications comprise a constant.

total working time is longer in the *AutonomyNoFeedback* condition, in which participants, on average, extend their working time by 1.19 units.[13]

Estimates from probit models confirm that participants in *AutonomyNoFeedback* are significantly more likely to buy additional working time than participants in the *AutonomyFeedback* condition.[14] The reported estimated marginal effects in column (1) of Table 1 indicate that subjects in *AutonomyNoFeedback* are 24 percentage points more likely to buy additional working time compared to subjects in *AutonomyFeedback*. This effect is robust to controlling for subjects' age, gender, and risk attitude (see column (2)).[15] Finally, in column (3), we additionally control for task productivity in the test round of the experiment (task-specific skills), the quantity of correctly solved Raven matrices (cognitive skills) as well as correctly submitted answers during the first two batches (i.e., ten pictures) of the working phase (task productivity). Again, the main result remains unaffected.[16]

Lastly, we note that results from further stages corroborate our main findings. We found a smaller and statistically not significant effect for stage two (marginal effect equal to 5.4 percentage points, p-value = 0.572) and a larger and statistically significant effect for stage three (marginal effect equal to 47 percentage points, p-value < 0.001).[17] Not least, we were able to replicate this study's main result in a companion project (Dohmen and Shvartsman, 2022) with 67 subjects.[18]

3.2. Subjective Stress

It is conceivable that working time autonomy reduces ex ante perceived stress because participants feel assured that they could work longer if they expect not to reach the performance target during the regular working time. Likewise, it appears plausible that performance feedback under working time autonomy further reduces perceived stress because participants would know whether or not they would reach the performance target within regular working time and hence whether they should buy additional working time. At the same time, we have documented that working time autonomy without

[13] Conditional on buying working time, there was no significant difference in the number of batches that participants bought in the two *Autonomy* conditions (two-sided t-test, p-value = 0.632).

[14] Coefficient estimates of these models are provided in Appendix A in Table A.3.

[15] These results are also robust to using an incentivized lottery choice measure for risk preference.

[16] Results from Poisson regressions confirm that subjects under *AutonomyNoFeedback* buy significantly (at the 5 percent level) more working time batches than subjects in *AutonomyFeedback*, albeit only at the 10 percent level (p-value = 0.060), once we include all controls. These results are available from the authors upon request.

[17] Note that if we pool the data from stage one and stage two, the results are similar to the results from stage one. These results are available from the authors upon request.

[18] The average incidence of purchase of extra working time across all four samples is 0.213 in the control group and 0.521 in the treatment group. If this reflected the true effect size, we would be powered at 80% (at alpha=0.05) to detect the effect with a sample size of 70 subjects. If we calculated the mean of the incidence of purchases of extra working time weighted for the number of observations in each sample the weighted average would be 0.206 in the control and 0.515 in the treatment group. Based on that effect size we would be powered with a sample size of 68 subjects. If we additionally neglected the second study, in which the bonus payment was higher, we would be powered with a sample size of 90 subjects. Note that these power analyses are based on ex post calculation of observed effect sizes.

Table 2
Subjective Stress.

	Primary Appraisal	Stress before task	Stress after task
Dependent variable	(1) (OLS)	(2) (OLS)	(3) (OLS)
AutonomyNoFeedback	-0.234	−0.409*	-0.395
	(0.186)	(0.228)	(0.246)
NoAutonomyFeedback	-0.083	0.032	0.033
	(0.187)	(0.229)	(0.247)
AutonomyFeedback	-0.070	-0.317	-0.146
	(0.192)	(0.236)	(0.254)
Constant	3.543***	2.707***	3.098***
	(0.136)	(0.167)	(0.180)
Observations	175	175	175
$Adj. - R^2$	-0.007	0.016	0.005

Notes: Condition *NoAutonomyNoFeedback* serves as reference category. The values in parentheses represent standard errors. */**/*** denotes statistical significance at the 10/5/1% level.

performance feedback leads to overexertion of effort. If subjects anticipate this as a stressful experience, they might perceive higher ex-ante stress under *Autonomy*. Since we have measures of subjective stress in all four treatment conditions, we can also assess the role of *Autonomy* and *Feedback* on perceived self-reported stress.

To this end, we regress the three subjective, i.e., self-reported, stress measures on indicator variables for the treatment conditions. The associated results are displayed in Table 2, where *NoAutonomyNoFeedback* is the reference category. Column (1) refers to results with respect to the primary appraisal index, column (2) to results with respect to stress after the announcement of the work conditions but before task onset, column (3) to post-task stress evaluation, i.e., at the end of the work phase.

As is visible from column (1), there are no significant differences between conditions with respect to the primary appraisal index.[19] For the analysis of perceived stress before task onset (column (2)), we do not find statistically significant differences for four of the six possible bilateral post-estimation comparisons.[20] The exceptions are the comparison between *NoAutonomyNoFeedback* vs. *AutonomyNoFeedback* at the 10% level (p-value = 0.075) and *NoAutonomyFeedback* vs. *AutonomyNoFeedback* at the 5% level (p-value = 0.048). Lastly, the analysis for post-task stress (compare column (3)) reveals that for the latter comparison, i.e., *NoAutonomyFeedback* vs. *AutonomyNoFeedback*, there is a statistical difference at the 10% level (p-value = 0.075). In contrast, all other bilateral comparisons do not reveal any statistically significant results.

These findings suggest that working time autonomy, if anything, mildly reduces perceived stress. Also, there is no evidence that performance uncertainty increases perceived stress in our setting. However, certainty about performance in a situation where there is no leeway in reacting to a (bad) performance feedback by extending one's working time appears to be perceived as moderately more stressful. Overall, it should be noted that we only inspected subjective stress measures and that it would be interesting to assess the impact of working conditions on objectively measured stress.

4. Concluding Discussion

Our analysis has revealed that a substantial fraction of subjects with working time autonomy in the *AutonomyNoFeedback* condition chose to incur a monetary cost for working longer. In our stylized laboratory setting, these subjects ended up providing more than the effort level necessary to secure the bonus, highlighting that uncertainty can induce overexertion of effort when subjects possess the possibility to work longer. Likewise, the results of our analysis, which contrast performance uncertainty with certainty, indicate that feedback provision can shield workers from the overprovision of effort.

Likely, overexertion of effort in our experiment results from motives aimed at reducing the probability of not achieving the performance target, which might stem from prudence or loss aversion. Subjects are willing to forgo earnings to increase their subjective probability of meeting the performance target and securing the bonus. The responses to our debriefing questionnaire are consistent with this conjecture. 70 percent of our subjects, who were initially assigned to the *Autonomy* conditions, flag the motive that the opportunity to buy into extra-working time increased their expectation to achieve the bonus, which is one of six motives that we suggested for buying extra working units.[21]

[19] We find qualitatively and quantitatively the same results for the sub-constructs challenge and threat indices and hence do not report these in the paper. The associated results are available from the authors upon request.

[20] There are no differences for *AutonomyNoFeedback* vs. *AutonomyFeedback* (p-value = 0.686), *AutonomyFeedback* vs. *NoAutonomyFeedback* (p-value = 0.130), *AutonomyFeedback* vs. *NoAutonomyNoFeedback* (p-value = 0.181), and *NoAutonomyNoFeedback* vs. *NoAutonomyFeedback* (p-value = 0.890).

[21] Participants could select up to three motives. The motives for buying extra time were: *(i)* reduces the time pressure (12.5%), *(ii)* grants autonomy (22.73%), *(iii)* increases flexibility (28.41%), *(iv)* reduces stress (35.23%), *(v)* increases expectation to reach the bonus (70.45%), and *(vi)* because the task is fun (11.36%). We administered this questionnaire after the entire experiment, i.e., after participants worked three times on the task.

The precautionary effort motive is an important mechanism that leads to overexertion of effort under working time autonomy in the absence of feedback but other behavioral biases, such as overconfidence, might counterbalance this effect. The relative importance of such countervailing factors depends on relative marginal costs of effort in regular working time vs. additional working time. As the marginal costs of working on the real effort task in our experiment was low during regular working time, the relative importance of factors that induce underprovision of effort were limited.[22]

Notably, our results suggest that working time autonomy mildly reduces perceived stress. Nonetheless, we cannot rule out a stress response to such work conditions in actual work settings. In our experiment, the potential overtime work is minimal, amounting only to a very short episode of a couple of minutes, and comes at a relatively low monetary cost.[23] In practice, these effort costs are arguably much higher.

While our design contrasts the extreme conditions *AutonomyFeedback* and *AutonomyNoFeedback*, which represent polar cases compared to a real-world environment, the associated insights may be particularly informative for managers. Our findings indicate that feedback can curb the overexertion of effort that arises in the presence of working time autonomy. In particular, precise feedback seems to be most conducive to increasing efficiency because subjects who are certain about their performance are less inclined to provide personally costly effort beyond a level necessary to secure a bonus. Therefore, high-quality feedback has a direct impact on reducing the effort costs of workers, which complements other potential beneficial effects of high-quality feedback on employment relationships such as increasing intrinsic motivation (see, e.g., Drouvelis and Paiardini, 2022).

One advantage of our experimental study is that we can observe the additional effort costs that participants incur regarding forgone earnings. A spontaneous objection could be that working time autonomy should increase workers' flexibility and reduce effort costs so that working time autonomy that comes with costly effort contradicts its very idea. However, working more is always associated with opportunity costs. We believe that our design feature of costly effort ensures a higher degree of external validity precisely because working more than necessary is costly in actual work settings, at least with respect to work-life balance.[24] In addition, though arguably less important, we incentivize subjects in our setting to abstain from slacking, as not turning to a problem in regular working time would come at a direct cost because subjects would have to compensate for this by providing personally costly overtime. We thereby refrain workers from procrastination and can therefore be assured that the provision of extra working time is not due to intertemporal effort allocation.[25]

We believe that our experimental design has several attractive features that can be exploited in other studies, pursuing to answer different questions pertaining to autonomy, performance certainty via feedback provision, effort provision, and incentives. Importantly, it enables researchers to manipulate working time autonomy and the conditions in which it is granted, control productivity, and vary effort costs. Despite these attractive features, we also recognize that our design entails limitations with respect to the scope and definition of working time autonomy. First, we model only a partial aspect of working time autonomy, and our experimental design can, therefore, not capture the effects of all associated aspects such as discretion regarding task order. Second, since laboratory sessions typically do not last longer than two hours there are limits to the analysis of effects of working time autonomy over more extended periods – for example, potential benefits of flexible working schedules such as optimal intertemporal effort provision.

Finally, due to the possibility of hedging we did not elicit beliefs about performance in the main work phase in an incentive-compatible manner.[26] We acknowledge, nonetheless, that beliefs would be informative for understanding effort provision in the absence of feedback. In particular, data on the subjective belief distribution of performance allows for measuring subjects' confidence about their performance and, thereby, subjects' perceived performance uncertainty. Hence, data on the distribution of subjects' beliefs could provide insights on the interaction of uncertainty and the intensity of the precautionary motive.

Our findings point to the importance of an attentive design of incentive structures and careful consideration of the conditions under which working time autonomy is granted to employees. While previous research has demonstrated various benefits entailed by autonomy (see, e.g., Nijp et al., 2012) – we also find mild support for a stress-reducing effect of autonomy – our findings suggest that under uncertainty, it can induce unnecessary effort costs. Hence, employers should be alerted that autonomy can harm workers in particular conditions.

[22] Non-incentivized questions about performance during total working time do not indicate widespread overconfidence. A large majority even underestimate their own performance.

[23] Over 94 percent of the subjects in the *Autonomy* conditions work three or less extra minutes, while two subjects buy into the maximum amount of nine extra minutes.

[24] Mas and Pallais (2017) estimate a relatively low willingness to pay for scheduling flexibility, which they partially explain by the fact that family members and friends desire concurrent schedules.

[25] It turns out that there is very little evidence for slacking off in our experiment. We tracked for each working unit whether subjects submitted an answer, which might also result from not answering in time: only 21 subjects did not submit an answer once, five subjects did not submit an answer twice, one further subject did not submit an answer three times, and two subjects finished working prematurely during regular working time.

[26] We only asked a non-incentivized question about performance (see, footnote [22]) to avoid the possibility of hedging. Alternatively, one could have, of course, randomize whether individuals are paid for actual performance or for the belief elicitation. However, such an approach would have reduced subjects' expected bonus payment and hence both their effort choices and potentially subjective stress.

Declaration of Competing Interest

none.

Data availability

The replication material for the study is available at DOI https://dx.doi.org/10.15185/izadp.16028.1

Appendix A

A.1. Additional Tables

Table A.1
Coefficients of Variation.

Task	*Authors*	*Coefficient of Variation*	*Mean*	*SD*	*N*	*Min*	*Max*	*Task Duration*	*Stage of Experiment*
Anagrams	Charness and Villeval (2009)	0.489	9.314	4.557	153	0	24	4 minutes	period four
Encryption Task	Erkal et al. (2011)	0.173	89.093	15.396	108	48	135	20 minutes	experiment one, stage one
Encryption Task	Kuhn and Villeval (2015)	0.104	55.546	5.756	174	38	68	4 minutes	part one
Multiplication Task	Dohmen and Falk (2011)	0.555	23.042	12.785	358	2	67	5 minutes	productivity stage
Pressing Keys	DellaVigna and Pope (2018)	0.345	1,936.329	668.218	9,861	1	3,950	10 minutes	main experiment
Sliders	Gill and Prowse (2012)	0.272	22.034	5.991	60	1	33	2 minutes	period one, first mover
Stuffing Letters	Carpenter et al. (2010)	0.299	10.505	3.138	224	0.4	21.6	30 minutes	main experiment (adj. perf.)
Summation Task	Weber and Schram (2017)	0.483	4.96	2.395	200	0	12	8 minutes	round one, employees
Dots Estimation Task	This Paper	0.106	37.726	4.016	175	11	44	9 minutes	stage one

Notes: N is the number of observations, SD is the standard deviation. The coefficient of variation was calculated using the mean and SD indicated in Gill and Prowse (2012, p. 482, Table 1). For all other publications, the publicly available data sets were used to calculate the coefficient of variation.

Table A.2
Balance Table.

	AutonomyNoFeedback	*NoAutonomyNoFeedback*	*NoAutonomyFeedback*	*AutonomyFeedback*	p-value
Age	24.021	25.659	23.935	24.585	0.439
	(5.302)	(7.098)	(4.814)	(4.254)	
Female (%)	0.596	0.659	0.674	0.610	0.848
	(0.496)	(0.480)	(0.474)	(0.494)	
Final Mathematics Grade	2.110	2.171	2.163	2.262	0.908
	(0.932)	(0.975)	(0.949)	(1.041)	
Productivity Test Round	0.770	0.824	0.787	0.815	0.400
	(0.161)	(0.162)	(0.176)	(0.164)	
Raven Matrix Test	4.830	4.976	5.109	5.244	0.788
	(1.810)	(1.943)	(2.292)	(1.841)	
Risk (1-11 scale)	5.745	5.585	5.565	5.561	0.977
	(2.617)	(2.302)	(2.007)	(2.169)	
Risk (MPL)	7.809	8.024	7.630	7.244	0.486
	(2.081)	(2.797)	(1.982)	(2.528)	
N	47	41	46	41	

Notes: N is the number of observations. The values in parentheses represent standard deviations. All p-values are from joint F-tests. */**/*** denotes statistical significance at the 10/5/1% level.

Table A.3
Working Time Purchase Decisions.

Dependent variable	Purchase of Additional Working Time		
	(1) (Probit)	(2) (Probit)	(3) (Probit)
AutonomyNoFeedback	0.626**	0.631**	0.732**
	(0.276)	(0.282)	(0.297)
Age		-0.014	-0.008
		(0.029)	(0.029)
Female		−0.510*	−0.485*
		(0.286)	(0.291)
Risk		0.055	0.053
		(0.059)	(0.062)
Productivity Test Round			0.042
			(0.940)
Raven Matrix Test			0.070
			(0.080)
Productivity Batch 1 and 2			-0.187
			(0.115)
Constant	−0.546***	-0.227	0.680
	(0.207)	(0.836)	(1.411)
Observations	88	88	88
Pseudo R^2	0.044	0.083	0.112

Notes: Condition *AutonomyFeedback* serves as reference category. The values in parentheses represent standard errors. */**/*** denotes statistical significance at the 10/5/1% level.

A.2. Chronological Sequence of the Study

After arrival at the laboratory, subjects were seated in private cubicles with computers. The welcome screen provided general information about the experiment (i.e., rules, payout, duration, …). Thereafter, participants solved a timed Raven matrix test. This was followed by a test phase of the real-effort task. In this test phase, subjects completed five units of the real effort task but did not receive feedback about their performance. Before proceeding, we administered the general and domain-specific risk (Dohmen et al., 2011; Richter et al., 2013) as well as the reciprocity questionnaires (Fehr and Schmidt, 2006; Perugini et al., 2003; Richter et al., 2013). Subjects then entered the work phase (including the elicitation of the primary appraisal index, pre-and post-task subjective stress measures, and other feelings), where instructions with respect to feedback and working time autonomy depended on the condition. After task completion, we also asked subjects to guess (non-incentivized) how many estimates they had submitted correctly in the *NoFeedback* conditions. Next, we also implemented two further stages, set up in the same way as stage one, to elicit sorting decisions into working time regimes. In stage two, subjects worked under the opposite working time regime compared to stage one, i.e., workers under *Autonomy* now faced fixed working time and vice versa. In the third stage, subjects could choose to work under *Autonomy* or under a fixed working time regime. In this paper, we focus on data from the first stage of the experiment as the first stage allows for a clean comparison of effort choice. At the beginning of stage two, we elicited risk via a multiple price list format. We also surveyed further personality traits, attitudes, and preferences for exploratory reasons, which we did not use in this study; these are listed in the Online Appendix. Moreover, all screens of the experiment, i.e., including screens of data not used in this study, are depicted in the Online Appendix.

At the end of each session, we surveyed subjects about the experiment. We first asked subjects whether they had used the option to buy into additional working time during the experiment and subsequently asked the reasons for their decision, where subjects entered the response into an open answer field. We then administered a questionnaire allowing subjects to choose up to three from six potential motives for buying additional working time. A demographics questionnaire concluded the experiment, before subjects received information on their payout.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jebo.2023.05.043

References

Akerlof, G.A., 1982. Labor contracts as partial gift exchange. The Quarterly Journal of Economics 97 (4), 543–569.
Avgoustaki, A., Bessa, I., 2019. Examining the link between flexible working arrangement bundles and employee work effort. Human Resource Management 58 (4), 431–449.
Avgoustaki, A., Frankort, H.T.W., 2019. Implications of work effort and discretion for employee well-being and career-related outcomes: An integrative assessment. ILR Review 72 (3), 636–661.
Azmat, G., Bagues, M., Cabrales, A., Iriberri, N., 2019. What you don't know... can't hurt you? A natural field experiment on relative performance feedback in higher education. Management Science 65 (8), 3714–3736.

Azmat, G., Iriberri, N., 2016. The provision of relative performance feedback: An analysis of performance and satisfaction. Journal of Economics & Management Strategy 25 (1), 77–110.
Beckmann, M., Cornelissen, T., Kräkel, M., 2017. Self-managed working time and employee effort: Theory and evidence. Journal of Economic Behavior & Organization 133, 285–302.
Bell, L.A., Freeman, R.B., 2001. The incentive for working hard: Explaining hours worked differences in the US and Germany. Labour Economics 8 (2), 181–202.
Blanes i Vidal, J., Nossol, M., 2011. Tournaments without prizes: Evidence from personnel records. Management Science 57 (10), 1721–1736.
Bock, O., Baetge, I., Nicklisch, A., 2014. hroot: Hamburg registration and organization online tool. European Economic Review 71, 117–120.
Carpenter, J., Matthews, P.H., Schirm, J., 2010. Tournaments and office politics: Evidence from a real effort experiment. American Economic Review 100 (1), 504–517.
Charness, G., Villeval, M.-C., 2009. Cooperation and competition in intergenerational experiments in the field and the laboratory. American Economic Review 99 (3), 956–978.
Corgnet, B., Hernán-González, R., 2019. Revisiting the trade-off between risk and incentives: The shocking effect of random shocks? Management Science 65 (3), 1096–1114.
Corgnet, B., Hernán-González, R., Schniter, E., 2015. Why real leisure really matters: Incentive effects on real effort in the laboratory. Experimental Economics 18 (2), 284–301.
DellaVigna, S., Pope, D., 2018. What motivates effort? Evidence and expert forecasts. The Review of Economic Studies 85 (2), 1029–1069.
Dohmen, T., Falk, A., 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. American Economic Review 101 (2), 556–590.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. Journal of the European Economic Association 9 (3), 522–550.
Dohmen, T., Rohde, I., Stolp, T., 2023. Tournament incentives affect perceived stress and hormonal stress responses. Experimental Economics Advance online publication.
Dohmen, T., & Shvartsman, E. (2022). Stress, incentives, and autonomy. Mimeo, University of Bonn.
Drouvelis, M., Paiardini, P., 2022. Feedback quality and performance in organisations. The Leadership Quarterly 33 (6), 101534.
Eeckhoudt, L., Huang, R.J., Tzeng, L.Y., 2012. Precautionary effort: A new look. Journal of Risk and Insurance 79 (2), 585–590.
Eriksson, T., Poulsen, A., Villeval, M.C., 2009. Feedback and incentives: Experimental evidence. Labour Economics 16 (6), 679–688.
Erkal, N., Gangadharan, L., Nikiforakis, N., 2011. Relative earnings and giving in a real-effort experiment. American Economic Review 101 (7), 3330–3348.
EWCS (2016). How are your working time arrangements set? (working time) visualisation : European bar chart by : Age, all - european working conditions survey - data visualisation EWCS 2016. https://www.eurofound.europa.eu/data/european-working-conditions-survey?locale=EN&dataSource=EWCS2017NW&media=png&width=740&question=y15_Q42&plot=euBars&countryGroup=linear&subset=agecat_3&subsetValue=All.
Fehr, E., Schmidt, K.M., 2006. The economics of fairness, reciprocity and altruism – Experimental evidence and new theories. Handbook of the Economics of Giving, Altruism and Reciprocity 615–691.
Fiesler, C. [@cfiesler]. (2021, June 9). *The problem with workload (and thus work-life balance) in academia is: You can always do more. (a thread based on a recent personal epiphany.)* [Tweet]. Twitter. https://twitter.com/cfiesler/status/1402634919516512259?s=03, last accessed: May 02, 2023.
Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171–178.
Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C.E., Falk, A., 2007. Social comparison affects reward-related brain activity in the human ventral striatum. Science 318 (5854), 1305–1308.
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 (6), 599–610.
Gibbs, M., 2017. How is new technology changing job design? IZA World of Labor.
Gill, D., Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. American Economic Review 102 (1), 469–503.
Goerg, S.J., Kube, S., Radbruch, J., 2019. The effectiveness of incentive schemes in the presence of implicit effort costs. Management Science 65 (9), 4063–4078.
Kelly, E.L., Moen, P., 2007. Rethinking the clockwork of work: Why schedule control may pay off at work and at home. Advances in Developing Human Resources 9 (4), 487–506.
Kuhn, P., Villeval, M.C., 2015. Are women more attracted to co-operation than men? The Economic Journal 125 (582), 115–140.
Lott, Y., 2014. Working time autonomy and time adequacy. What if performance is all that counts? Discussion Paper. Institute of Economic and Social Research (WSI).
Mas, A., Pallais, A., 2017. Valuing alternative work arrangements. American Economic Review 107 (12), 3722–3759.
Nijp, H.H., Beckers, D.G.J., Geurts, S.A.E., Tucker, P., Kompier, M.A.J., 2012. Systematic review on the association between employee worktime control and work-non-work balance, health and well-being, and job-related outcomes. Scandinavian Journal of Work, Environment & Health 38 (4), 299–313.
Perugini, M., Gallucci, M., Presaghi, F., Ercolani, A.P., 2003. The personal norm of reciprocity. European Journal of Personality 17 (4), 251–283.
Prendergast, C., 2002. The tenuous trade-off between risk and incentives. Journal of Political Economy 110 (5), 1071–1102.
Richter, D., Metzing, M., Weinhardt, M., Schupp, J., 2013. SOEP scales manual. SOEP Survey Papers. DIW Berlin.
Simon, H.A., 1955. A behavioral model of rational choice. The Quarterly Journal of Economics 69 (1), 99–118.
Simon, H.A., 1972. Theories of bounded rationality. In: McGuire, C.B., Radner, R. (Eds.), Decision and organization. Amsterdam: North-Holland, pp. 161–176.
Sloof, R., Van Praag, C.M., 2010. The effect of noise in a performance measure on work motivation: A real effort laboratory experiment. Labour Economics 17 (5), 751–765.
Weber, M., Schram, A., 2017. The non-equivalence of labour market taxes: A real-effort experiment. The Economic Journal 127 (604), 2187–2215.