Competition and Fatigue At Work

Vera Angelova (TU Berlin)
Thomas Giebe (Linnaeus University)
Radosveta Ivanova-Stenzel (TU Berlin)

Discussion Paper No. 134

December 20, 2018

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Competition and Fatigue at Work*

Vera Angelova[a], Thomas Giebe[b,*], Radosveta Ivanova-Stenzel[a]

[a] *Technische Universität Berlin, Faculty of Economics and Management, Str. des 17. Juni 135, 10623 Berlin, Germany*
[b] *Centre for Discrimination and Integration Studies, Linnaeus University, Sweden, thomas.giebe@lnu.se*

Abstract

We study theoretically and experimentally the role of fatigue and recovery within a competitive work environment. At work, agents usually make their effort choice in response to competition and monetary incentives. At the same time, they have to take into account fatigue, which accumulates over time if there is insufficient recovery. We model a sequence of work periods as tournaments that are linked through fatigue spillovers, inducing a non-time-separable decision problem. We also allow for variations in incentives in one work period, in order to analyze spillover effects to the work periods "before" and "after". Making recovery harder should, generally, reduce effort. This theoretical prediction is supported by the experimental data. A short-term increase in incentives in one period should lead to higher effort in that period, and, due to fatigue, to strategic resting before and after. Our experimental results confirm the former, whereas we do not find sufficient evidence for the latter. Even in the presence of fatigue, total effort should positively respond to higher-powered incentives. This is not supported by our data. Removing fatigue, we find the expected increase in total effort. For work environments, this may imply that the link between monetary incentives and effort provision becomes weaker in the presence of fatigue or insufficient recovery between work periods.

Keywords: fatigue, recovery, incentives, experiment, tournament
JEL classifications: C72, C91, D9, J22, J33, M5, M52

1. Introduction

Designing incentive schemes for dynamic competitive work environments requires understanding of how incentives and variations in incentives over time affect employees' performance. It has been established that an increase in monetary incentives positively affects effort provision in the short run (Jenkins Jr et al., 1998). Over time, however, individuals often respond with strategic effort allocation, which can have unintended consequences (Asch, 1990, Miklós-Thal and Ullrich, 2015). Indeed, a few studies have found that an increase in incentives in dynamic environments might not result in an increase in performance, even though there is a clear behavioral response (Goette and Huffman, 2006, Angelova et al., 2018).

*Corresponding author
*We thank Paul Schweinzer in particular and are grateful for discussions with Florian Englmaier, Dorothea Kübler, Martin Siegel, participants at ESA 2018 in Berlin, Contests: Theory and Evidence 2018 in Norwich, The Lisbon Meetings in Game Theory and Applications 2018, the Annual Meeting of Swedish Economists 2018, the UCSD-Rady Visitors Conference on Incentives and Behavior Change in Modica 2017, the department seminars at Université Laval and Linnaeus University. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 "Rationality and Competition" as well as support by the Berlin Centre for Consumer Policies (BCCP) are gratefully acknowledged.

A natural explanation for why observed performance in dynamic settings might not respond to changes in incentives as expected is the presence of fatigue and the need for recovery. The relevance of fatigue in work environments is obvious. Fatigue is an established empirical phenomenon (Kant et al., 2003).[2] Recent empirical work shows that a large part of the working population accumulates fatigue during the work week, and (partially) recovers over the weekend (Åkerstedt et al., 2018). There is a small number of theoretical contributions that focus on effort choice in the presence of fatigue and recovery (Dragone, 2009, Baucells and Zhao, 2018). However, far too little is known about the role of fatigue and recovery in competitive work environments. Our study contributes to filling this gap.

One inherent feature of competition at work is fluctuations in incentives over time. These can be natural, or deliberately introduced as well as objective or perceived. In firms, competition for promotion as well as rewards based on relative performance are the prevalent form of workplace organization (Prendergast, 1999, Lazear, 1999). For example, promotions are understood as prizes in a competition between workers that takes place over a longer period (Lazear and Rosen, 1981). During that time, it is natural that work periods differ in their relevance towards promotion. For instance, an employee might have an important presentation on Tuesday, and regular office days for the rest of the week. Then the employee's performance on Tuesday will be more visible or count more towards promotion, or might be perceived as being more important than regular working days.

As an example for deliberate variations in incentives, consider sales contests. The business literature recommends varying the severity of competition over time by introducing short-term sales contests on top of existing incentive schemes (Roberge, 2015).

Following the organizational literature, we analyze work environments through the lens of tournament theory (Prendergast, 1999, Connelly et al., 2014). The starting point for our theoretical and experimental analysis is the rank-order tournament model of Lazear and Rosen (1981).[3] This (static) model has been widely used in the literature on management, organization, personnel economics, and experimental economics. We augment it in two dimensions.

First, we allow for a dynamic competitive environment, i.e., a *sequence* of tournaments. A sequence of tournaments can be interpreted as consecutive work periods. These periods can be working days, weeks, or months, and employees (partially) recover from fatigue over night, on weekends, or during holidays.[4] Second, we introduce fatigue and recovery through a *non-time-separable* cost function, in which marginal effort cost increases in previous periods' effort, and decreases with resting.

The model allows us to vary the quality of recovery between work periods. Depending on the choice of the fatigue parameter, a given effort choice leads to different degrees of fatigue. In addition, we allow for variations in incentives in a representative "middle" period, in order to analyze spillover effects to the work periods "before" and "after". A short-term increase in incentives in one period may justify higher effort in that period, but also strategic resting before and after, as fatigue requires a sensible effort allocation over time.

[2] Similarly, understanding fatigue is important in professional sports (Montgomery et al., 2008).

[3] See Dechenaux et al. (2015) for an overview of standard tournament models and corresponding experimental studies.

[4] For example, consider competition for promotion. In each work period, employees choose effort in order to contribute to being positively evaluated by their supervisors, relative to their co-workers. Monetary incentives in the form of tournament prizes can be understood as contributions ("points") towards positive performance evaluation and a favorable promotion decision in the (distant) future.

In our model, two agents interact in a sequence of three rank-order tournaments that are linked through the fatigue and recovery technology.[5] Agents need to choose an effort profile for these three periods, taking into account the monetary incentives profile and the non-time-separable effects of fatigue. The benchmark symmetric equilibrium exhibits a V-shaped effort profile under constant incentives over time and an inverse V-shape if monetary incentives in the middle tournament are sufficiently larger than in the tournaments before and after. Higher incentives in the middle tournament imply higher effort in that tournament. This is combined with strategic resting in the tournaments before and after. Overall, total effort, i.e., the sum of efforts in all tournaments, increases. In contrast, for a given incentive scheme, if recovery is made harder, i.e., if fatigue becomes more severe, agents respond with lower effort in all tournaments resulting in a decrease in total effort.

In the experiment, we vary subjects' ability to recover between tournaments. In line with comparative statics predictions, subjects choose lower effort in all tournaments if there is less opportunity to recover between tournaments. In addition, we employ two tournament incentive profiles, one with constant incentives over time, and one with higher-powered incentives in the middle tournament. Corresponding to theory, subjects choose higher effort in the middle tournament under the latter incentive scheme, whereas we observe insufficient strategic resting before and after. Regardless of the severity of fatigue, *total* effort does not significantly respond to changes in the incentive scheme. However, when we completely remove fatigue, total effort does increase as predicted, suggesting that the link between monetary incentives and effort provision is not as strong as expected when recovery is not complete.

To the best of our knowledge, this study is the first attempt to theoretically and experimentally investigate how the interplay between fatigue, recovery and the severity of competition affects effort provision in competitive work environments.[6]

2. A Tournament Model with Fatigue and Recovery

We develop a simple dynamic model with fatigue and recovery based on the seminal static rank-order tournament model of Lazear and Rosen (1981). In our model, two agents $i \in \{1, 2\}$ simultaneously choose unobservable costly effort e_{ti} in a sequence of three tournaments $t \in \{1, 2, 3\}$. In each tournament, agents compete for a winner prize W_t, while the loser receives the prize L_t. The winner of tournament t is the agent with the higher output, where output is the sum of effort e_{ti} and an i.i.d. productivity shock ϵ_{ti}. The probability that agent i wins tournament t against agent j is

$$\Pr(e_{ti} + \epsilon_{ti} > e_{tj} + \epsilon_{tj}) = \Pr(\epsilon_{tj} - \epsilon_{ti} < e_{ti} - e_{tj}) =: G(e_{ti} - e_{tj}), \tag{1}$$

where G is defined as the corresponding c.d.f. of the difference of the productivity shocks.

Introducing fatigue and recovery into this dynamic environment intuitively implies that effort choice in one tournament affects utility in that tournament as well as future tournaments.

Thus, we introduce a non-time-separable effort cost function as follows. Agent i's cost function in tournament t, C_{ti}, is a function of the present and all previous tournaments' effort choices. Fa-

[5]The model can easily be extended to a longer time-horizon and other incentive profiles.
[6]There are a few tournament models that explicitly deal with fatigue alone, e.g., Ryvkin (2011).

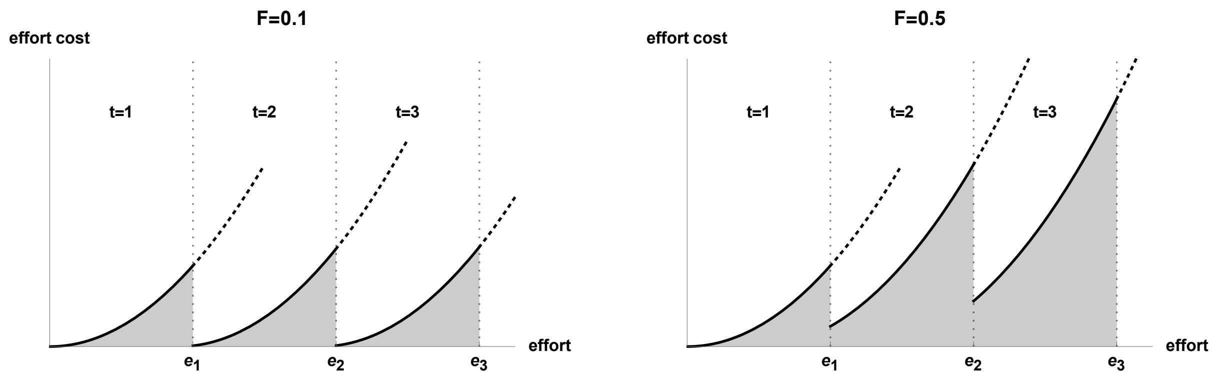


Figure 1: Illustration of the cost function for fatigue parameters $F = 0.1$, $F = 0.5$.

tigue is implemented as an increase in marginal cost that is caused by effort in previous tournaments, $\partial^2 C_{ti}/(\partial e_{ti}\partial e_{si}) > 0$ for all $s < t$. Recovery is modelled, first, through the size of that marginal-cost effect. Second, the contribution of a given tournament to future fatigue decays over time, i.e., $\partial^2 C_{ti}/(\partial e_{ti}\partial e_{si})$ is decreasing in the distance $t - s$.

The cost function is defined as

$$C_{ti} = k\left(\sum_{s=1}^{t} F^{t-s} e_{si}\right)^2 = \begin{cases} k(e_{1i})^2 & t = 1, \\ k(Fe_{1i} + e_{2i})^2 & t = 2, \\ k(F^2 e_{1i} + Fe_{2i} + e_{3i})^2 & t = 3, \end{cases} \tag{2}$$

with $F \in (0, 1)$ and $k > 0$.[7] Note the distinction between an agent's *fatigue level* or tiredness, represented by the value of the effort cost function at any point in time, and the *fatigue parameter* F, which represents the ability to recover between working periods. Intuitively, F can be seen as the quality of recovery between working days, e.g., a shorter night corresponds to a larger value of F.

Figure 1 illustrates the cost function, with effort choice (time spent working) on the horizontal, and effort cost (tiredness) on the vertical axis. It shows how an agent becomes more tired *during* working days, and (partially) recovers *between* working days. A low value of $F = 0.1$ (left panel) implies nearly full recovery between periods, whereas $F = 0.5$ (right panel) illustrates accumulating fatigue, for example, during the work week.[8]

Denote the prize spread, i.e., the additional payoff of a winner, by $P_t = W_t - L_t$. We consider prize spreads of the form $(P_1, P_2, P_3) = (P, P + \delta, P)$ with $\delta \geq 0$. Comparing the cases $\delta = 0$ and $\delta > 0$ allows us to study how the severity of competition affects effort provision including spillovers between periods. The latter are caused by an agent's response to fatigue and recovery.

Strategically, our model corresponds to a one-shot game where each agent's action is represented by

[7] This cost function satisfies the above described properties. In particular, for $F \in (0, 1)$, $0 < 2kF^2 = \partial^2 C_{3i}/(\partial e_{3i}\partial e_{1i}) < 2kF = \partial^2 C_{2i}/(\partial e_{2i}\partial e_{1i}) = \partial^2 C_{3i}/(\partial e_{3i}\partial e_{2i})$. Note that we do not include $F = 0$ and $F = 1$ in the definition of the cost function, to simplify some arguments. For $F = 0$, the cost function collapses to a standard quadratic effort cost function for the standard model without fatigue, i.e., three independent tournaments. $F = 1$ is the extreme and irrelevant case of no recovery at all. The parameter k can be varied to select suitable parameters for experimental tests.

[8] Figure 1 plots equal efforts in all tournaments, regardless of optimality.

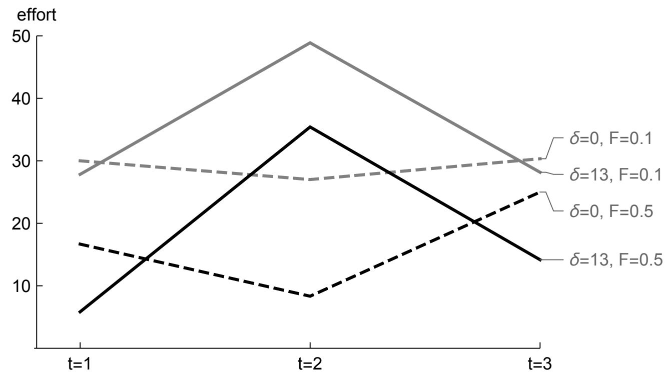


Figure 2: Equilibrium effort profiles

a three-dimensional effort vector (profile).[9] Thus, each risk-neutral agent i maximizes (total) expected payoff from participation in the three tournaments,

$$\sum_{t=1}^{3} \left(G(e_{ti} - e_{tj})P_t + L_t - C_{ti}\right). \tag{3}$$

Given the symmetric setup of the game and following the literature on rank-order tournaments, we restrict attention to symmetric pure-strategy equilibria, $e^*_{ti} = e^*_{tj} = e^*_t$. This implies that $G'(e^*_{ti} - e^*_{tj}) = G'(0)$ is a constant (as in a standard rank-order tournament).

In any symmetric pure-strategy equilibrium with positive efforts,

$$e^*_1 = \frac{G'(0)}{2k}\left(P(1-F) - F\delta\right) \tag{4}$$

$$e^*_2 = \frac{G'(0)}{2k}\left(P(1-F)^2 + \delta(1+F^2)\right) \tag{5}$$

$$e^*_3 = \frac{G'(0)}{2k}\left(P(1-F+F^2) - F\delta\right). \tag{6}$$

As can be seen in (4) and (6), positive equilibrium efforts cease to exist for sufficiently large F for given $\delta > 0$. In order to reasonably test the theoretical predictions, we chose the model parameters for our experiment such that predicted efforts satisfy three criteria: (i) they constitute a unique pure-strategy equilibrium, (ii) they are positive, in order to differentiate between a dropout decision in the experiment and a zero effort decision that corresponds to the prediction, (iii) they are sufficiently different between and within our experimental treatments.[10] Table 1 reports the equilibrium predictions and Figure 2 plots the equilibrium effort profiles for the model parameters that have been used in the experiment.

In the following, we discuss several comparative statics predictions regarding the variation in incentives, δ, and the fatigue parameter, F. Based on (4) – (6), we can make some immediate observations. A larger prize spread P unambiguously increases effort in all periods. Increasing the prize spread in tournament 2 only, $\delta > 0$, leads to higher effort in tournament 2, and lower effort in tournaments 1 and 3. The latter can be interpreted as strategic resting before and after a period of higher incentives. A simple computation

[9] A rational agent has no reason to revise the chosen effort profile between tournaments.
[10] The proof of uniqueness is available on request.

confirms that the effect on total effort is unambiguously positive.

Proposition 1 (Comparative Statics: Incentives)**.** *Let* (e_1^*, e_2^*, e_3^*) *characterize a symmetric pure-strategy equilibrium with positive efforts. Then, ceteris paribus, an increase in the prize spread from* P *to* $P + \delta$ $(\delta > 0)$ *in tournament 2 implies*

- *higher effort in tournament 2,*
- *lower effort in tournaments 1 and 3 (strategic resting),*
- *higher total effort* $e_1^* + e_2^* + e_3^*$*.*

The proof of this and the following proposition can be found in the Supplementary Material. Let us keep the fatigue parameter F fixed in the following discussion of Proposition 1. Start with the case $\delta = 0$ (flat incentives over time). In symmetric pure-strategy equilibrium with positive efforts, the effort profile is V-shaped ($e_1^* > e_2^* < e_3^*$, see Figure 2): Effort in tournament 3 is highest, because this effort does not imply spillovers into the future. In that sense, it is the cheapest effort and, correspondingly, the other two efforts are smaller in order to induce low fatigue at the beginning of tournament 3. Effort in tournament 1 is the second highest, because at the beginning of tournament 3, fatigue from tournament 1 has decayed to a larger extent than that of tournament 2. Finally, effort in tournament 2 is the smallest, as that effort has the strongest fatigue spillover on tournament 3.

Now compare the flat incentive scheme ($\delta = 0$) to one with higher incentives in the second tournament ($\delta > 0$, see Figure 2). As a direct response to the larger prize spread, optimal effort in tournament 2 is higher. However, due to fatigue spillovers, there need to be (optimal) adjustments in the other two tournaments as well. The effort in tournament 1 needs to be reduced in order to start tournament 2 at a lower fatigue level. The effort in tournament 3 will be reduced as well, because fatigue is high at the end of tournament 2, implying high effort cost in tournament 3. We interpret these two latter effects as strategic resting.

Our next result deals with the impact of the fatigue parameter F on effort levels. It specifies the conditions under which a larger fatigue parameter unambiguously decreases effort in all periods.

Proposition 2 (Comparative Statics: Fatigue)**.** *Let* (e_1^*, e_2^*, e_3^*) *characterize a symmetric pure-strategy equilibrium with positive efforts. Then, ceteris paribus, increasing the fatigue parameter from* F *to* $F+\Delta F$ *leads to*

- *lower efforts in tournament 1, in tournament 2 if* $2F+\Delta F < \frac{2P}{P+\delta}$*, and in tournament 3 if* $2F+\Delta F < \frac{P+\delta}{P}$*,*
- *lower total effort* $e_1^* + e_2^* + e_3^*$*.*

Proposition 2 shows that the intuitive result of lower effort under higher fatigue holds for moderate or small fatigue parameters. Restricting attention to such a parameter range is required if we want to consider fatigue levels that typically accumulate during the work week and from which employees recover over the weekend (Åkerstedt et al., 2018, see also Figure 1). For the experimental tests, we have chosen parameters for which the conditions in Proposition 2 are satisfied by a wide margin.

3. Experimental Design and Hypotheses

The design of the computerized laboratory experiment follows the set-up of the model.[11] In each round, two subjects simultaneously chose effort for each of three tournaments.[12] Efforts were chosen from the interval $[0,70]$, the random number was drawn from a uniform distribution with support $[-30,30]$. The loser prize was $L=30$. We implemented incentive profiles $(P, P+\delta, P)$, with prize spread $P=20$ in tournaments 1 and 3 in all treatments. We employed a 2 x 2 factorial design with $\delta \in \{0,13\}$ and fatigue parameter $F \in \{0.1, 0.5\}$ (see Table 1 for an overview of the four treatments together with the equilibrium predictions). "Hump" denotes the incentive scheme with $\delta=13$, i.e., the larger winner prize in the second tournament, whereas "Flat" refers to $\delta=0$. The abbreviations "L" and "H" refer to the low ($F=0.1$), respectively the high value ($F=0.5$) of the fatigue parameter. Subjects were randomly assigned to treatments, and each subject participated in one treatment only.

Each treatment consisted of 30 rounds. In each round, subjects interacted with the same opponent. Between rounds, subjects were randomly rematched within a matching group of 8. Per treatment, we had 6 matching groups which qualify as independent observations. Altogether, 192 subjects participated in the experiment, i.e., 48 subjects per treatment.

At the end of each round, subjects were informed about their own effort choice, cost, random number and output, as well as which prize they won and their total payoff, for each of the three tournaments. There was no feedback between individual tournaments, in order to control for feedback effects (e.g. discouragement). Three rounds were randomly selected and paid in cash at the end of the experiment.

At the end of the experiment, we also collected data on cognitive ability (using a 5-minute Raven test), individual risk preferences (using a lottery experiment), and some demographic characteristics. The experiment was conducted with students (37% female), mostly from economics, natural sciences, or engineering at Technische Universität Berlin in the spring and summer of 2017.

	Low fatigue, $F=0.1$	**High** fatigue, $F=0.5$
Flat incentives, $\delta=0$	**FlatL**	**FlatH**
(e_1^*, e_2^*, e_3^*)	$(30.0, 27.0, 30.3)$	$(16.7, 8.3, 25.0)$
Hump-shaped incentives, $\delta=13$	**HumpL**	**HumpH**
(e_1^*, e_2^*, e_3^*)	$(27.8, 48.9, 28.2)$	$(5.8, 35.4, 14.2)$

Table 1: Treatments and equilibrium predictions

Our experimental analysis focusses on the comparative statics predictions stated in Propositions 1 and 2. This leads to the following two hypotheses.

Hypothesis 1 (Incentives)**.** *For a given fatigue parameter F,*

*(i) effort in the middle tournament is higher in the **Hump**-treatments than in the **Flat**-treatments,*

[11] For the instructions see the Supplementary Material.

[12] To explain the cost function, we provided a graph and a cost calculator on each decision screen in addition to the verbal explanation and the mathematical formula in the instructions. We verified that subjects understood the rules of the game with a quiz.

*(ii) efforts in the tournaments before and after are lower in the **Hump**-treatments than in the **Flat**-treatments, and,*

*(iii) total effort is higher in the **Hump**-treatments than in the **Flat**-treatments.*

Hypothesis 2 (Fatigue)**.** *For a given incentive parameter δ, effort in each tournament and, hence, total effort, are higher in the **L**-treatments than in the **H**-treatments.*

Our last hypothesis provides the test of the model's point predictions (see (4) – (6)).

Hypothesis 3. *The efforts in each treatment are equal to the theoretically predicted efforts.*

4. Results

Table 2 gives an overview of average effort across treatments as well as total effort, whereas Figure 3 depicts average effort by tournament and treatment over time together with the theoretical predictions. In the Hump-treatments, average effort in tournament 2 is always above effort in the other two tournaments, and also above effort in tournament 2 in the respective Flat-treatments. For a given fatigue parameter, effort levels in tournaments 1 and 3 in the Hump-treatments are slightly lower than those in the Flat-treatments. Regardless of the incentive scheme, efforts in all tournaments are lower in the H-treatments compared to the L-treatments. Efforts in the L-treatments, except for tournament 2 in HumpL, are close to the theoretical benchmark. In all treatments, we observe a downward adjustment of effort over time with behavior stabilizing in the second half of the experiment.

Treatment	**Effort**			
	e_1	e_2	e_3	$e_1 + e_2 + e_3$
FlatL	31.71 (21.71)	29.64 (20.73)	31.79 (19.81)	93.14 (54.91)
HumpL	29.58 (20.57)	39.00 (21.80)	28.37 (19.22)	96.95 (54.46)
FlatH	21.94 (21.64)	19.57 (17.47)	23.91 (18.90)	65.42 (45.44)
HumpH	18.33 (18.94)	27.89 (21.92)	21.13 (18.27)	67.34 (39.99)
Additional:				
Flat0	29.87 (18.23)	28.65 (17.89)	29.25 (18.07)	87.77 (49.29)
Hump0	31.35 (20.74)	43.15 (22.14)	32.50 (21.54)	107.00 (56.39)

Table 2: Average effort, standard deviation in parentheses

A multivariate regression analysis combined with post-estimation tests confirms that for any given fatigue parameter, (i) subjects exert significantly more effort in tournament 2 in the Hump-treatments than in the Flat-treatments (FlatL vs. HumpL, resp. FlatH vs. HumpH, $p \leq 0.011$); (ii) in tournaments

1 and 3, effort in the Hump-treatments does not significantly differ from effort in the Flat-treatments.[13] Hence, although we observe a tendency for strategic resting, it is not statistically significant. In all tournaments, subjects provide significantly less effort when resting is made harder (FlatL vs. FlatH, resp. HumpL vs. HumpH, $p \leq 0.0175$). We obtain the same results with Wilcoxon-Mann-Whitney tests.[14]

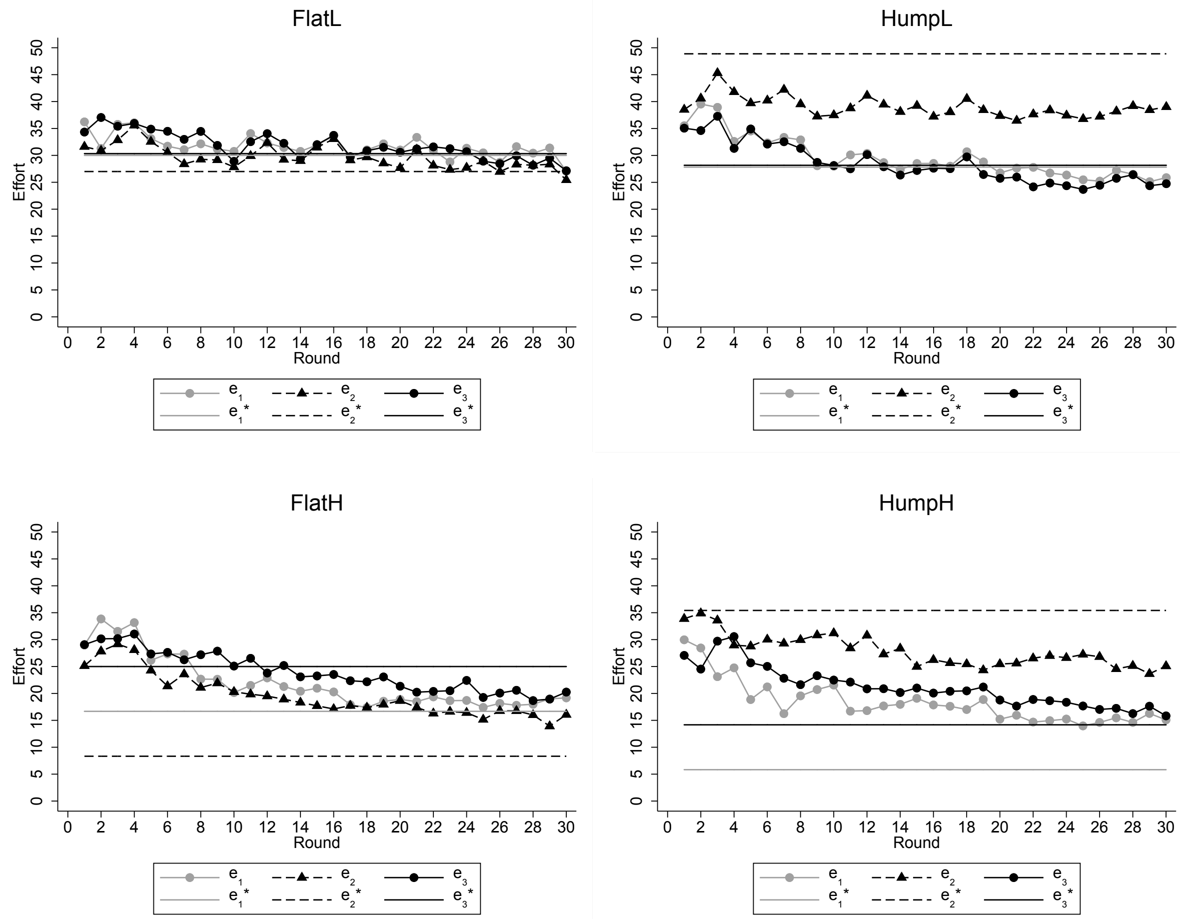


Figure 3: Average effort and equilibrium prediction by tournament over time

The comparison of the average efforts in the three tournaments to the point predictions provides mixed evidence. In the L-treatments, i.e., in the case of nearly full recovery, average effort differs from the theoretical prediction only in tournament 2 in HumpL ($p = 0.0021$, Wilcoxon-Mann-Whitney test). In the H-treatments, behavior is never in line with the theoretical prediction except for tournament 3 in FlatH.

All results are confirmed if we consider the second half of the experiment only. Details on the statistical analysis can be found in the Supplementary Material.

Result 1.

- *Compared to a flat incentive scheme, stronger incentives in the middle tournament lead to higher*

[13]The effort choices for tournaments 1, 2, and 3 are dependent on each other.

[14]All non-parametric tests are based on averages on the level of statistically independent matching groups. Tests are always two-sided.

effort in that tournament, supporting H1(i). In the tournaments before and after, there is no significant strategic resting, rejecting H1(ii).

- *For a given incentive scheme, effort in all tournaments is lower when resting is made harder, supporting H2.*
- *Overall, the theory rationalizes observed behavior better when there is nearly full recovery or when the incentives remain constant over time, partially supporting H3.*

Dependent variable:	(1)	(2)	(3)
Total effort	All rounds	All rounds	Second half
Reference category	FlatL	FlatL	FlatL
FlatH	-27.72***	-26.19***	-32.24***
	(8.225)	(8.197)	(9.303)
HumpH	-25.80***	-24.32***	-28.05***
	(7.937)	(8.052)	(8.722)
HumpL	3.810	5.605	2.953
	(9.190)	(8.776)	(9.812)
Constant	93.14***	107.1***	81.51**
	(6.451)	(29.53)	(32.80)
Controls	No	Yes	Yes
Observations	5,760	5,760	2,880
Number of subjects	192	192	192
$\beta_{FlatH} = \beta_{HumpH}$	0.7806	0.7990	0.6176
$\beta_{HumpL} = \beta_{HumpH}$	0.0002	0.0002	0.0006

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

Table 3: GLS regression results (robust standard errors in parentheses); post-estimation tests (last two rows)

As shown in Table 2 (last column), total effort in the L-treatments is substantially higher than in the H-treatments. In contrast, total effort in the Hump-treatments is only minimally above total effort in the corresponding Flat-treatments. Table 3 displays the results of linear regressions comparing total effort across treatments. All specifications are generalized least squares (GLS) models with random effects at the subject level and clustered standard errors at the matching group level to account for correlated decisions by the same subject and within the same matching group. Specification (1) contains treatment dummies whose coefficients can either directly be compared to the reference category FlatL, or via post-estimation tests (p-values are reported in the last two rows of Table 3). Specification (2) additionally contains several control variables.[15] Specification (3) is identical to (2) with one exception: the analysis in (3) is conducted on the second half of the experiment only, i.e. when behavior has stabilized.[16]

Specification (1) shows that total effort in FlatH is significantly lower than in FlatL ($p = 0.001$). Similarly, total effort in HumpH is below total effort in HumpL ($p = 0.0002$). However, total effort does not significantly differ between FlatL and HumpL ($p = 0.678$) as well as between FlatH and HumpH

[15]The control variables include gender, cognitive ability, risk-aversion, self-reported impulsiveness, age, and the number of semesters studied. These variables are not significant.

[16]In all treatments, the correlation between total effort and round becomes insignificant in the second half of the experiment.

($p = 0.7806$). These results do not change when we include controls (specification (2)), or run the analysis on the second half of the experiment (specification (3)).

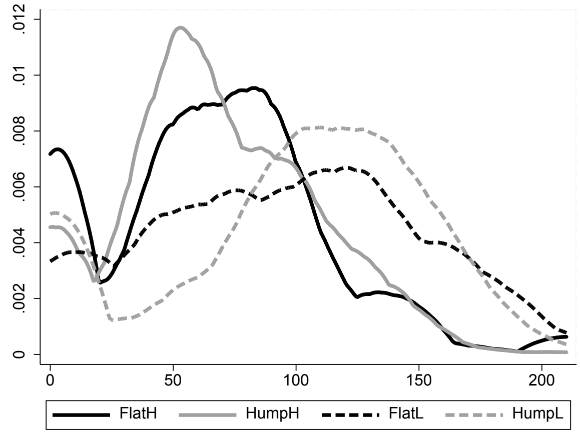


Figure 4: Density of total effort by treatment

Figure 4 compares the distributions of total effort in the different treatments. It clearly shows that treatments are grouped in pairs according to the fatigue parameter, with both H-treatments being shifted to the left. We compare the distributions of individual average total effort across treatments with a Kolmogorov-Smirnov test. Indeed, there is no significant difference between the Hump-treatments and the Flat-treatments for any fatigue parameter ($p \geq 0.368$), whereas distributions differ significantly for a given incentive scheme between the L- and H-treatments ($p \leq 0.002$).

Result 2.

- *Introducing higher-powered incentives in one tournament does not lead to higher total effort. Thus, we reject H1(iii).*
- *When resting is made harder, total effort decreases. This result is consistent with H2.*

In our experiment, the effort profiles should be V-shaped in the Flat-treatments and inverse V-shaped in the Hump-treatments (see Figure 2). For each subject, we computed the average effort profile for the second half of the experiment and classified it into one of three categories, see Table 4. In all treatments, the most frequent shape of the effort profiles is, indeed, the predicted one. Subjects whose effort profiles are classified as "Other" seem to respond mainly to fatigue. They concentrate effort in either the first or the last tournament.

Effort profile	HumpH	HumpL	FlatH	FlatL
V-shaped ($e_1 \geq e_2 \leq e_3$)	27%	23%	**54%**	**54%**
Inverse-V-shaped ($e_1 \leq e_2 > e_3$ or $e_1 < e_2 = e_3$)	**44%**	**65%**	15%	12%
Other ($e_1 < e_2 < e_3$ or $e_1 > e_2 > e_3$)	29%	13%	31%	33%

Table 4: Relative frequency of subjects' average effort profile (second half of experiment)

Recall that, given the model parameters for our experiment, we should observe strictly positive efforts in all tournaments. In fact, only 38% of subjects choose positive efforts throughout the experiment, whereas the rest (62%) drops out in at least one tournament.[17] Among the latter group, 63% in the H-treatments and 37% in the L-treatments drop out in more than half of the rounds. Nobody drops out in all tournaments and all rounds. We find some evidence for 'compensating behavior', in particular in the H-treatments: For instance, subjects who drop out in tournaments 1 and 3 in HumpH exert 50% more effort in tournament 2 than subjects who always choose positive efforts.[18] In all treatments, subjects who drop out exert significantly less total effort than those who never drop out. However, the shapes of the observed effort profiles are similar between the two groups.

One important finding so far is that total effort does not positively respond to higher-powered incentives. To check whether this and all other experimental results depend on the presence of fatigue, we conducted two additional treatments without fatigue, Flat0 and Hump0. We also used them to check for feedback effects. The additional treatments are identical to those investigated so far with two exceptions: (i) there is no fatigue, i.e., $F = 0$, and (ii) in half of the 12 matching groups, subjects received the feedback after each tournament rather than after each sequence of three tournaments. Theoretically, these feedback differences are irrelevant, which was confirmed by the data. Thus, we will not discuss them any further.

The last two rows of Table 2 provide the descriptives for the two treatments without fatigue. As predicted by the theory, subjects exert significantly more effort in tournament 2 in Hump0 than in Flat0 ($p = 0.000$), though the increase is insufficient compared to the theoretical benchmark ($p = 0.0021$).[19] Strategic resting is not an issue in the absence of fatigue. Indeed, there is no significant difference in observed average effort in tournaments 1 and 3 between Hump0 and Flat0 ($p = 0.645$, resp. $p = 0.319$). Moreover, in line with the theory, the switch from the Flat to the Hump incentive scheme leads to a significant increase in total effort ($p = 0.021$).[20] Our results suggest that it is due to the presence of fatigue that total effort is not higher in the Hump-treatments than in the Flat-treatments.

5. Concluding Remarks

Many work environments are characterized by the simultaneous presence of fatigue, recovery, different incentive schemes and competition. We propose a simple theoretical model that combines all these elements, based on the well-known rank-order tournament model of Lazear and Rosen (1981). We test the model's comparative statics predictions in a controlled environment. We find that subjects strategically respond to variations in the fatigue technology and the different incentive schemes. In particular, they reduce effort if recovery is made harder. A one-time increase in incentives induces a positive short-term response in effort and only a slight tendency for strategic resting in the periods before and after. However, we do not observe an increase in total effort. In contrast, when fatigue is removed we do observe that higher-powered incentives lead to higher total effort. Our results suggest that the conventional belief that high(er) incentives induce high(er) effort does not necessarily hold in the presence of fatigue. This

[17] Female subjects are more likely to exert positive efforts throughout the experiment.

[18] See the Supplementary Material for details.

[19] The theoretical prediction for this case follow directly if we set $F = 0$ in our model, corresponding to three strategically independent static tournaments.

[20] The results for the additional treatments are based on the same statistical analysis as reported so far. Details can be found in the Supplementary Material.

implies that fatigue needs to be taken into account when designing incentive schemes in competitive work environments.

Overall, our neoclassical model, in which fatigue and recovery are considered as part of an agent's production technology, cannot fully explain the observed behavior. A promising path of future research is to identify the potential behavioral biases or decision heuristics that might be responsible for the observed behavior. For example, the pattern of observed insufficient effort in the tournament with higher-powered incentives as well as insufficient resting before and after (as compared to the theoretical predictions) is reminiscent of the anchoring and insufficient adjustment heuristic also known as *pull-to-center bias* and first observed in a laboratory experiment on the newsvendor problem (Schweitzer and Cachon, 2000). Another avenue might involve a different way of modeling the fatigue and recovery technology.

References

Åkerstedt, T., Ghilotti, F., Grotta, A., Zhao, H., Adami, H.O., Trolle-Lagerros, Y., Bellocco, R., 2018. Sleep duration and mortality — does weekend sleep matter? Journal of Sleep Research (e12712). doi:10.1111/jsr.12712.

Angelova, V., Giebe, T., Ivanova-Stenzel, R., 2018. Does a short-term increase in incentives boost performance? Economics Letters 166, 31–34. doi:10.1016/j.econlet.2018.02.004.

Asch, B.J., 1990. Do incentives matter? The case of Navy recruiters. Industrial & Labor Relations Review 43, 89S–106S. doi:10.1177/001979399004300306.

Baucells, M., Zhao, L., 2018. It is time to get some rest. Management Science (forthcoming). doi:10.1287/mnsc.2017.3018.

Connelly, B.L., Tihanyi, L., Crook, T.R., Gangloff, K.A., 2014. Tournament theory: Thirty years of contests and competitions. Journal of Management 40, 16–47. doi:10.1177/0149206313498902.

Dechenaux, E., Kovenock, D., Sheremeta, R.M., 2015. A survey of experimental research on contests, all-pay auctions and tournaments. Experimental Economics 18, 609–669. doi:10.1007/s10683-014-9421-0.

Dragone, D., 2009. I am getting tired: Effort and fatigue in intertemporal decision-making. Journal of Economic Psychology 30, 552–562. doi:10.1016/j.joep.2009.03.008.

Goette, L., Huffman, D., 2006. Incentives and the allocation of effort over time: The joint role of affective and cognitive decision making. IZA DP No. 2400 .

Jenkins Jr, G.D., Mitra, A., Gupta, N., Shaw, J.D., 1998. Are financial incentives related to performance? A meta-analytic review of empirical research. Journal of Applied Psychology 83, 777–787. doi:10.1037/0021-9010.83.5.777.

Kant, I., Bültmann, U., Schröer, K., Beurskens, A., Van Amelsvoort, L., Swaen, G., 2003. An epidemiological approach to study fatigue in the working population: the Maastricht cohort study. Occupational and Environmental Medicine 60, i32–i39. doi:10.1136/oem.60.suppl_1.i32.

Lazear, E.P., 1999. Personnel economics: Past lessons and future directions presidential address to the Society of Labor Economists, San Francisco, May 1, 1998. Journal of Labor Economics 17, 199–236. doi:10.1086/209918.

Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. The Journal of Political Economy 89, 841–864. doi:10.1086/261010.

Miklós-Thal, J., Ullrich, H., 2015. Career prospects and effort incentives: Evidence from professional soccer. Management Science 62, 1645–1667. doi:10.1287/mnsc.2015.2211.

Montgomery, P.G., Pyne, D.B., Hopkins, W.G., Dorman, J.C., Cook, K., Minahan, C.L., 2008. The effect of recovery strategies on physical performance and cumulative fatigue in competitive basketball. Journal of Sports Sciences 26, 1135–1145. doi:10.1080/02640410802104912.

Prendergast, C., 1999. The provision of incentives in firms. Journal of Economic Literature 37, 7–63. doi:10.1257/jel.37.1.7.

Roberge, M., 2015. The right way to use compensation. Harvard Business Review 93, 70–75.

Ryvkin, D., 2011. Fatigue in dynamic tournaments. Journal of Economics & Management Strategy 20, 1011–1041. doi:10.1111/j.1530-9134.2011.00314.x.

Schweitzer, M.E., Cachon, G.P., 2000. Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. Management Science 46, 404–420. doi:10.1287/mnsc.46.3.404.12070.

Competition and Fatigue at Work
SUPPLEMENTARY MATERIAL

Vera Angelova[a], Thomas Giebe[b,*], Radosveta Ivanova-Stenzel[a]

[a] *Technische Universität Berlin, Mikroökonomie, Sekr. H91, Str. des 17. Juni 135, 10623 Berlin, Germany*
[b] *Centre for Discrimination and Integration Studies, Linnaeus University, Sweden*

December 15, 2018

Contents

*Corresponding author, thomas.giebe@lnu.se

1. Theory

1.1. The distribution of noise terms

The distribution of the difference between the two players' noise terms is denoted by

$$\Pr(e_{ti}+\epsilon_{ti}>e_{tj}+\epsilon_{tj})=\Pr(\epsilon_{tj}-\epsilon_{ti}<e_{ti}-e_{tj})=:G(e_{ti}-e_{tj}). \tag{1}$$

Assuming that each noise term ϵ_{ti} is independently drawn from a continuous uniform distribution with support $[-a,a]$, $a\in\mathbb{R}^+$, the density $G'(x)$ is

$$G'(x)=\begin{cases}\frac{x}{4a^2}+\frac{1}{2a} & -2a\le x\le 0\\ -\frac{x}{4a^2}+\frac{1}{2a} & 0<x\le 2a\\ 0 & \text{otherwise.}\end{cases} \tag{2}$$

Therefore, $G'(0)=\frac{1}{2a}$. In order to evaluate deviations from the equilibrium candidate, the corresponding distribution $G(x)$ is needed:

$$G(x)=\int_{-\infty}^{x}G'(\tilde{x})d\tilde{x}=\begin{cases}0 & x<-2a\\ \frac{x^2}{8a^2}+\frac{x}{2a}+\frac{1}{2} & -2a\le x\le 0\\ -\frac{x^2}{8a^2}+\frac{x}{2a}+\frac{1}{2} & 0<x\le 2a\\ 1 & x>2a.\end{cases} \tag{3}$$

1.2. Proof of Propositions 1 and 2

Proof of Proposition 1. The equilibrium efforts are linear functions of δ. Compute the derivatives with respect to δ,

$$\frac{\partial e_1^*}{\partial\delta}=-F\frac{G'(0)}{2k}<0, \tag{4}$$

$$\frac{\partial e_2^*}{\partial\delta}=(1+F^2)\frac{G'(0)}{2k}>0, \tag{5}$$

$$\frac{\partial e_3^*}{\partial\delta}=-F\frac{G'(0)}{2k}<0. \tag{6}$$

Adding these marginal effects, we find

$$\frac{\partial(e_1^*+e_2^*+e_3^*)}{\partial\delta}=(1-F)^2\frac{G'(0)}{2k}>0. \tag{7}$$

□

Proof of Proposition 2. Effort in tournament 1 is strictly decreasing in F for any $F\in(0,1)$, P and $\delta\ge 0$, $\frac{\partial e_1^*}{\partial F}=-(P+\delta)\frac{G'(0)}{2k}<0$.

Regarding effort in tournament 2 (resp., tournament 3), consider an increase in the fatigue parameter from F to $F+\Delta F$, and denote the corresponding equilibrium efforts in tournaments 2 (resp., tournament 3) by e_2^* and e_2^{**} (resp., e_3^* and e_3^{**}).

Straightforward computation shows that the inequality $e_2^{**} < e_2^*$ is equivalent to $2F + \Delta F < \frac{2P}{P+\delta}$.

A similar result is obtained for effort in tournament 3, where the inequality $e_3^{**} < e_3^*$ is equivalent to $2F + \Delta F < \frac{P+\delta}{P}$.

Total effort is strictly decreasing in F for any $F \in (0,1)$, P, and $\delta \geq 0$

$$\frac{\partial(e_1^* + e_2^* + e_3^*)}{\partial F} = (F-1)(2P+\delta)\frac{G'(0)}{2k} < 0. \tag{8}$$

$\square$

2. Results

With "Hump" we denote the incentive scheme with $\delta = 13$ (i.e., a larger winner prize in the second tournament), whereas "Flat" refers to $\delta = 0$. The abbreviations "L" and "H" refer to the low ($F = 0.1$), respectively the high value ($F = 0.5$) of the fatigue parameter.

2.1. Descriptives

Table 1 lists average effort levels and the equilibrium predictions across tournaments and treatments for the whole experiment and for its second half.

Figure 1 visualizes the averages across tournaments and treatments together with the theoretical predictions. While average effort is very similar across tournaments in a given Flat-treatment, in the Hump-treatments, we observe the expected increase in average effort in tournament 2, as well as the drop in effort in tournaments 1 and 3 compared to the Flat-treatments. Furthermore, subjects seem to correctly react to the increased fatigue parameter: on average they exert less effort in treatments with high fatigue parameter than in treatments with low fatigue parameter for a given incentive scheme. Finally, in the Hump-treatments, the observed reaction to the increased prize spread in tournament 2 is much less pronounced than theory predicts.

Treatment		**Effort** e_1	e_2	e_3
FlatL	Whole experiment	31.71 (21.71)	29.64 (20.73)	31.79 (19.81)
	Second half	30.77 (21.12)	28.58 (20.01)	30.22 (19.78)
	Nash	30.00	27.00	30.33
HumpL	Whole experiment	29.58 (20.57)	39.00 (21.80)	28.37 (19.22)
	Second half	27.11 (20.53)	38.04 (22.18)	25.73 (18.91)
	Nash	27.83	48.88	28.17
FlatH	Whole experiment	21.94 (21.64)	19.57 (17.47)	23.91 (18.90)
	Second half	18.51 (20.23)	16.68 (16.53)	20.91 (18.69)
	Nash	16.67	8.33	25.00
HumpH	Whole experiment	18.33 (18.94)	27.89 (21.92)	21.13 (18.27)
	Second half	15.82 (17.21)	25.69 (22.01)	18.40 (17.32)
	Nash	5.83	35.42	14.17

Table 1: Average effort with standard deviation in parentheses and theoretical prediction (Nash)

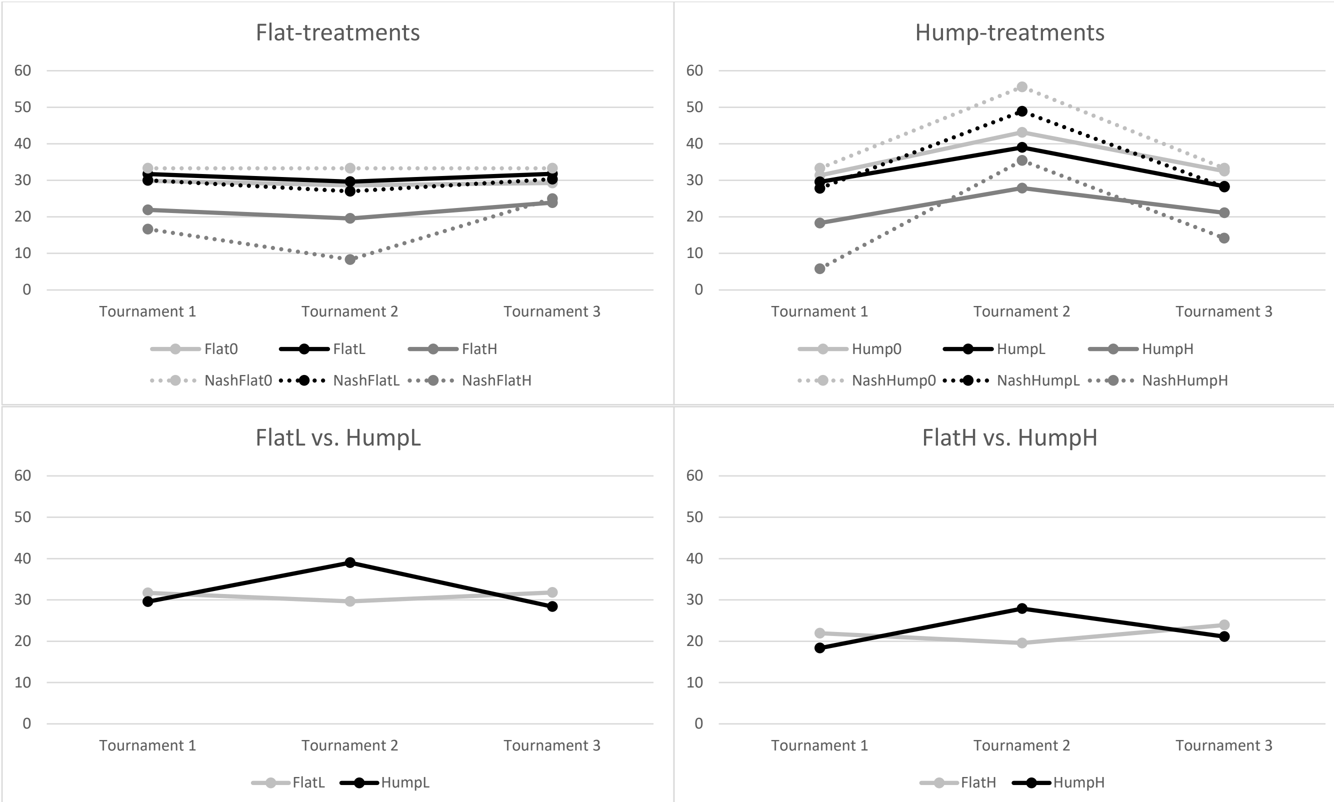


Figure 1: Average effort by tournament and treatment together with Nash predictions (upper two diagrams)

2.2. Testing the Hypotheses

To test our hypotheses, we need to compare effort in a given tournament across treatments. As the effort choices for tournaments 1, 2, and 3 are dependent on each other, we simultaneously regress efforts in tournament 1, tournament 2, and tournament 3 on dummy variables for the four treatments using multivariate regression analysis. Based on this regression, we run post-estimation tests that pairwise compare the estimated coefficients of the (appropriate) dummy variables for each effort. The unit of observation is average effort for each subject and each tournament. The regression results and the results of the post-estimation tests are reported in Table 2 and Table 3, respectively. The conclusions from Table 2 and Table 3 do not change when considering the second half of the experiment only.

We also tested our hypotheses using Wilcoxon-Mann-Whitney tests. Table 4 provides the two-sided p-values for the whole sample and for the second half. The tests confirm our findings from the regression analysis. We find statistical support for H2 regarding fatigue. As for H1, regarding incentives, we only find partial support: effort in tournament 2 in the Hump-treatments is significantly higher than in the Flat-treatments. We find some evidence for strategic resting but only with the low fatigue parameter and only with experienced subjects (second half of the experiment). Contrary to H1(ii), total effort in the Hump-treatments is not higher than in the Flat-treatments.

Table 5 shows the p-values from the comparison of efforts in tournaments 1, 2, and 3 to the point predictions for the whole experiment and for its second half. In FlatL, there is no significant difference between behavior and the predictions. In HumpL, subjects choose efforts according to the predictions in tournaments 1 and 3 but not in tournament 2. In FlatH, effort choices in tournaments 1 and 2 significantly differ from the predictions and in HumpH this is the case in all tournaments. These results do not change substantially when we perform the same analysis on the second half of the experiment.

	e_1	e_2	e_3
FlatL-dummy	31.71***	29.64***	31.79***
	(2.372)	(2.387)	(2.136)
HumpL-dummy	29.58***	39.00***	28.37***
	(2.372)	(2.387)	(2.136)
FlatH-dummy	21.94***	19.57***	23.91***
	(2.372)	(2.387)	(2.136)
HumpH-dummy	18.33***	27.89***	21.13***
	(2.372)	(2.387)	(2.136)
Observations	192	192	192
R-squared	0.718	0.769	0.768

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2: Multivariate regression comparing effort in the single tournaments across treatments (whole experiment)

	FlatL vs. HumpL	FlatH vs. HumpH	FlatL vs. FlatH	HumpL vs. HumpH
e_1	H1(ii): 0.5271	H1(ii): 0.2828	H2: 0.0040	H2: 0.0010
e_2	H1(i): 0.0061	H1(i): 0.0147	H2: 0.0032	H2: 0.0012
e_3	H1(ii): 0.2585	H1(ii): 0.3580	H2: 0.0098	H2: 0.0175

Table 3: Two-sided p-values from post-estimation tests comparing the regression coefficients in Table 2

	Whole experiment		Second half	
Hypothesis	For a given fatigue parameter			
Compare Flat to Hump	L	H	L	H
H1(i): e_1	0.3367	0.2623	0.1495	0.5218
H1(i): e_2	0.0374	0.0104	0.0163	0.0104
H1(i): e_3	0.2002	0.5218	0.0374	0.7488
H1(ii): $e_1 + e_2 + e_3$	0.4233	0.6310	0.6310	0.6310
Hypothesis	For a given incentive scheme			
Compare L to H	Flat	Hump	Flat	Hump
H2: e_1	0.0039	0.0065	0.0065	0.0163
H2: e_2	0.0065	0.0104	0.0039	0.0104
H2: e_3	0.0250	0.0547	0.0250	0.0163
H2: $e_1 + e_2 + e_3$	0.0039	0.0039	0.0039	0.0065

Table 4: Two-sided p-values from Wilcoxon-Mann-Whitney tests based on six independent observations per treatment

Treatment	Observed vs. Nash		
	Whole experiment/Second half		
	e_1	e_2	e_3
FlatL	1.0000/0.3051	1.0000/1.0000	0.3051/0.3051
HumpL	0.3051/1.0000	0.0021/0.0021	1.0000/0.3051
FlatH	0.0403/0.3051	0.0021/0.0021	0.3051/0.3051
HumpH	0.0021/0.0021	0.0403/0.0403	0.0403/0.0403

Table 5: Two-sided $p-$values from Wilcoxon-Mann-Whitney Tests: Observed vs. Nash

2.3. Dropout Behavior

Dropout behavior is more prevalent in the H-treatments than in the L-treatments. The number of effort profiles with at least one zero effort is as follows: (FlatH, HumpH, FlatL, HumpL)=(639, 638, 335, 274). In fact, 73 subjects choose positive efforts in all tournaments throughout the experiment, whereas 119 subjects drop out in at least one tournament. Figure 2 plots on the vertical axis the number of subjects who choose an effort profile with at least one zero as often as indicated on the horizontal axis. In the L-treatments (upper two diagrams) the zero bars are clearly higher, indicating that in those treatments fewer subjects choose zero effort. In the H-treatments, we see more and also higher bars on the right-hand side of the distributions, showing that in those treatments, more subjects dropped out more often. We compare the four distributions with Kolmogorov-Smirnov tests. The tests reject the equality of distributions between FlatL and FlatH ($p = 0.018$), as well as between HumpL and HumpH ($p = 0.018$), while the equality of distributions is supported for FlatL vs. HumpL ($p = 0.960$), and FlatH vs. HumpH ($p = 0.997$). We conclude that drop out behavior is driven by the fatigue parameter rather than the incentive scheme.

Do subjects who drop out compensate a zero effort in one tournament by exerting more effort in the other tournaments? To answer this question we proceed as follows.

First, we identify all effort profiles with either

· zero effort in one tournament combined with positive efforts in the remaining two tournaments (column "one zero-effort" in Table 6)
  or

· zero effort in two tournaments combined with positive effort in the remaining tournament (column "two-zero efforts" in Table 6).

For both categories (and separately for each of the respective tournaments), we compute the average effort(s). Then, we use those numbers to construct a measure of "compensation" as a fraction of the average effort in the same tournament (1, 2, or 3) chosen by the subjects who never drop out. For instance, for all effort profiles with $(e_{1d}, e_{2d}, e_{3d}) = (0, 0, > 0)$, we are interested in the average $\overline{e_{3d}}$. We compute the ratio $\overline{e_{3d}}/\overline{e_{3n}}$, where $\overline{e_{3n}}$ is the average effort in tournament 3 of those subjects who always choose positive effort profiles, i.e., $(e_{1n}, e_{2n}, e_{3n}) = (> 0, > 0, > 0)$, throughout the experiment.

Table 6 reports the measure for all tournaments and treatments. Note that "compensating" behavior is present if that measure is larger than 1. We find some evidence for such behavior, in particular in the H-treatments. For instance, in FlatH zero efforts in tournaments 1 and 2 are compensated by 20% more effort in tournament 3 on average (see column "$e_1 = e_2 = 0$" in Table 6). In HumpH, zero efforts

	One zero-effort						Two zero-efforts		
	$e_1 = 0$		$e_2 = 0$		$e_3 = 0$		$e_1 = e_2 = 0$	$e_2 = e_3 = 0$	$e_1 = e_3 = 0$
	e_2	e_3	e_1	e_3	e_1	e_2	e_3	e_1	e_2
FlatL	0.6	0.8	0.9	0.9	0.6	0.7	1.1	1.0	1.0
HumpL	1.4	0.8	1.2	1.3	1.1	1.1	no obs.	no obs.	1.0
FlatH	0.9	1.3	1.1	1.2	1.2	1.2	1.2	1.0	1.1
HumpH	1.5	0.9	1.2	1.3	0.7	1.4	1.4	1.0	1.5

Table 6: "Compensating behavior" of subjects who drop out in at least one tournament

in tournaments 1 and 3 are compensated by 50% more effort in tournament 2 on average (see column "$e_1 = e_3 = 0$" in Table 6).

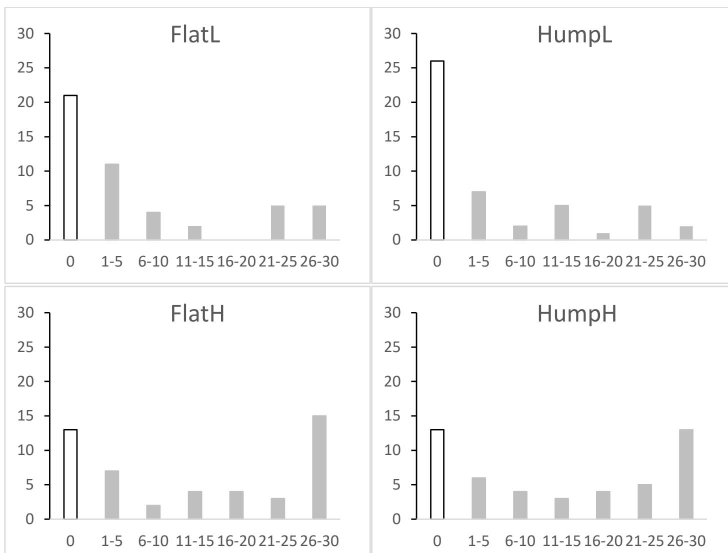


Figure 2: Distribution of subjects' dropout frequency

2.4. Additional Treatments: Flat0 and Hump0

Table 7 provides a descriptive overview of the choices of subjects in the additional treatments Flat0 and Hump0, for the whole experiment, and for its second half, together with the theoretical predictions (Nash).

Table 8 shows the results of multivariate regressions comparing effort choices across Flat0 and FlatH for each tournament. The reference category is Flat0. The nonsignificant coefficients of the Hump0-dummy for tournaments 1 and 3 point out that there is no difference in tournament 1 (respectively 3) between Flat0 and Hump0. The highly significant Hump0-coefficient for e_2 indicates that effort in tournament 2 in Hump0 is significantly above effort in tournament 2 in Flat0. Results do not change if the analysis is conducted on the second half of the experiment only or with linear random effects regressions with clustered standard errors on the matching group level that we run separately for effort in each tournament.

We compare effort choices in the additional treatments to the theoretical predictions and provide the p-values in Table 9. Only in tournaments 1 and 3 of Hump0 are choices in line with theory. Again, results do not change when considering the second half of the experiment only.

Table 10 provides the results of regressions that compare total effort in Flat0 to total effort in Hump0. Specification (2) adds all controls to specification (1), while specification (3) performs the analysis of specification (2) on the second half of the experiment. The Hump0-dummy is always positive and significant, indicating that total effort in Hump0 is higher than in Flat0.

Table 11 and Table 12 provide evidence that decisions with feedback do not differ from decisions without feedback and that this is true for both, Flat0 and Hump0.

Treatment		**Effort**		
		e_1	e_2	e_3
Flat0	Whole experiment	29.87(18.23)	28.65(17.89)	29.25(18.07)
	Second half	28.60(17.34)	27.81(16.47)	28.64(17.30)
	Nash	33.33	33.33	33.33
Hump0	Whole experiment	31.35(20.74)	43.15(22.14)	32.50(21.54)
	Second half	29.94(20.83)	41.54(23.16)	31.27(21.98)
	Nash	33.3	55.56	33.3

Table 7: Average effort with standard deviation in parentheses and theoretical prediction (Nash)

	e_1	e_2	e_3
Reference category	Flat0	Flat0	Flat0
Hump0-dummy	1.485	14.50***	3.249
	(3.210)	(3.065)	(3.244)
Constant	29.87***	28.65***	29.25***
	(2.270)	(2.167)	(2.294)
Observations	96	96	96
R-squared	0.002	0.192	0.011

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Multivariate regression, Flat0 vs. Hump0 (whole experiment)

Treatment	**Observed vs. Nash**		
	Whole experiment/Second half		
	e_1	e_2	e_3
Flat0	0.0403/0.0021	0.0403/0.0021	0.0403/0.0403
Hump0	0.3051/0.3051	0.0021/0.0021	1.0000/0.3051

Table 9: Two-sided $p-$values from Wilcoxon-Mann-Whitney Tests: Observed vs. Nash

	All rounds	All rounds+controls	Second half+controls
Hump0	19.23**	19.74***	18.18**
	(8.352)	(7.262)	(7.082)
Ability		3.698***	4.473***
		(1.291)	(1.459)
Risk aversion		-1.532	-1.314
		(2.803)	(3.093)
Impulsiveness		3.883**	4.083**
		(1.933)	(2.015)
Female		10.36	11.97
		(11.91)	(12.16)
Age		0.804	0.909
		(0.817)	(1.019)
Semesters		-0.774	-1.601
		(1.658)	(2.043)
Constant	87.77***	16.62	3.452
	(4.387)	(25.32)	(25.51)
Observations	2,880	2,880	1,440
Subjects	96	96	96

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: GLS regressions with random effects at the subject level and clustered standard errors at the matching group level

	e_1	e_2	e_3
Feedback-dummy	-2.693	-2.739	-4.022
	(3.984)	(3.792)	(3.957)
Constant	31.21***	30.02***	31.26***
	(2.817)	(2.681)	(2.798)
Observations	48	48	48
R-squared	0.010	0.011	0.022

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Flat0: feedback vs. no feedback, multivariate regression, whole experiment

3. Experiment

3.1. Parameters

In the experiment, we employed a 2 x 2 factorial design with $\delta \in \{0, 13\}$ and $F \in \{0.1, 0.5\}$. We set $k = 0.005$, $L = 30$, $P = 20$.[1] For the distribution of i.i.d. individual noise terms, ϵ_{ti}, we used a uniform distribution with support $[-30, 30]$. This implies $G'(0) = \frac{1}{60}$.

[1]When selecting suitable parameters for the experiment, we used a program written in Mathematica that helped us verify the equilibrium by numerically checking the deviation utility of every feasible deviation for all treatments on a grid that was much finer than that used in the experiment.

	e1	e2	e3
Feedback-dummy	-4.457	-6.757	-4.301
	(5.065)	(4.782)	(5.158)
Constant	33.58***	46.53***	34.65***
	(3.582)	(3.381)	(3.647)
Observations	48	48	48
R-squared	0.017	0.042	0.015

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Hump0: feedback vs. no feedback, multivariate regression, whole experiment

In the experiment, subjects were restricted to choose efforts for each tournament from $[0, 70]$, i.e., a fine grid of 71 distinct effort levels (including zero), comprising a total of 71^3 feasible effort vectors.

3.2. Procedures

The experiment was conducted at Technische Universität Berlin in the spring and summer of 2017. Most participants were students with a major in economics, natural sciences, or engineering. The experiment was programmed in z-Tree (Fischbacher, 2007). Subjects were invited to the experiment with ORSEE (Greiner, 2015). The experiment lasted around 90 minutes; average earnings were 25.44 EUR including a show-up fee of 5 EUR. Subjects were informed that the experiment would consist of three parts, and that instructions for each part would be distributed after the previous part was finished. In the instructions we used neutral language, referring to, e.g. 'effort' as 'number', 'tournament' as 'interaction', etc.

Part I was the heart of our experiment: there subjects made their effort choices in the different treatments. The instructions for Part I are provided in the next subsection.

In Part II we measured cognitive ability with a 5-minute Raven test. We recorded the number of correctly solved matrices for each participant. A higher number indicates higher ability. We paid 0.3 EUR per correctly solved matrix.

In Part III, we elicited individual risk preferences with a lottery experiment similar to Holt and Laury (2002). Participants were required to choose between two lotteries, A and B. Each lottery had two possible payoffs. The payoffs for lottery A were 2 EUR and 1.60 EUR, the payoffs for lottery B were 3.85 EUR and 0.10 EUR. The high payoffs in both lotteries were realized with the same probability p. Participants faced the choice between A and B ten times. From one choice to the next, the probability p increased from 10% to 100% in steps of 10%. At the end, one of the ten lottery pairs was randomly selected, played, and paid out. The control variable "Risk-aversion" used in the GLS regressions counts how many times a given subject chose lottery A (which was the safe choice), i.e., it takes values from 0 to 10, with higher numbers corresponding to higher risk-aversion.

At the end of each session we collected some additional data such as gender, age, the number of semesters studied and self-assessed impulsiveness (on a scale between 0 (not at all) to 10 (very impulsive)).[2]

[2]We asked subjects: "How do you judge yourself: Are you generally a person who contemplates for a long time before

3.3. Instructions

The following instructions were translated from German. They are identical across treatments except for the treatment variables. Below, we provide one set of instructions, where we indicate the differences between treatments with cursive font surrounded by brackets. The instructions for the additional treatments are available upon request.

Welcome to the experiment and thank you for participating!
General information

The instructions are the same for all participants.

Please read these instructions carefully. If there is something you do not understand, please raise your hand. We will then come to you and answer your questions privately.

You will make your decisions at the computer.

All decisions will remain anonymous. This means, that you will not know the identity of the other participants, and no participant will know your identity.

For simplification, the instructions are given in the masculine form.

The experiment consists of three parts. At the beginning of each part, you will receive detailed instructions. The parts are independent, i.e. your decisions in one part will not affect the results in any other part.

In every part of the experiment you will earn money. How exactly you will earn money will be described in the instructions.

Your earnings in this experiment (i.e. the sum of your earnings from all three parts) will be paid to you privately and in cash at the end of the experiment.

You will receive 5 EUR for showing up on time.

<u>Part 1</u>

All monetary amounts are expressed in the fictitious currency "ECU" (Experimental Currency Unit).

Part 1 consists of 30 rounds.

At the beginning of each round, groups of two participants will be formed. These groups of two will be randomly reassembled at the beginning of each subsequent round.

Each round consists of three interactions. For each interaction, every participant will make one decision. These three decisions will be made on the same screen.

In each interaction, the task of every participant will be to select an integer between 0 and 70. This number will cause costs expressed in ECU. The costs will increase exponentially in the selected number. A graph depicting this relationship between the selected number and its costs will be displayed on each decision screen (see the figure of the decision screen at the end of these instructions). The costs in interaction 1 will depend on the number selected in interaction 1 (number1). The costs in interaction 2 will depend on number1 and on the number selected in interaction 2 (number2); the costs in interaction 3 will

acting, i.e. who is not impulsive at all? Or, are you a person who acts without thinking too long, i.e. who is very impulsive?".

depend on number1, number2, and the number selected in interaction 3 (number3). The number selected in interaction 1 (number1) will affect the costs in interaction 2 more than the costs in interaction 3. On your decision screen, you will always be able to have calculated how the costs depend on the selected numbers in the three interactions. The costs will be calculated as follows:

Costs in interaction 1 $= 0.005 * (\text{number1})^2$
[*L-treatments:*] Costs in interaction 2 $= 0.005 * (0.1 * \text{number1} + \text{number2})^2$
[*H-treatments:*] Costs in interaction 2 $= 0.005 * (0.5 * \text{number1} + \text{number2})^2$
[*L-treatments:*] Costs in interaction 3 $= 0.005 * (0.01 * \text{number1} + 0.1 * \text{number2} + \text{number3})^2$
[*H-treatments:*] Costs in interaction 3 $= 0.005 * (0.25 * \text{number1} + 0.5 * \text{number2} + \text{number3})^2$

An illustration of the decision screen with explanations can be found at the end of these instructions.

For each interaction and for each participant, a real random number between -30 and +30 will be drawn independently and with equal probability.

For each interaction, the **result** of a participant is the sum of the number he selects and the random number drawn for him in this interaction:

result = selected number + random number

For each interaction, your result will be compared to the result of the other participant in your group of two. This will result in the following **amounts of money**:

	Interaction 1	Interaction 2	Interaction 3
The participant with the **higher** result will receive	50 ECU	[*Flat:*] 50 ECU [*Hump:*] 63 ECU	50 ECU
The participant with the **lower** result will receive	30 ECU	30 ECU	30 ECU

For each interaction, the costs of the selected number will be subtracted from this amount of money, leading to the following payoff for each interaction:

Interaction payoff (ECU) = amount of money (ECU) − costs for the selected number (ECU)

At the end of each round, you will see the following information for each interaction: your selected number, your random number, your result, your amount of money, the costs for your selected number and your interaction payoff.

At the end of the experiment, **three** rounds from Part 1 will be randomly selected. Your earnings from Part 1 will be equal to the sum of your interaction payoffs in the randomly selected rounds. Please note that your decisions in every round are important, as each round may potentially affect your earnings from Part 1.

Please note that a negative interaction payoff is possible. You will get a negative interaction payoff, if the costs for the selected number exceed the amount of money in this interaction.

The exchange rate is: **20 ECU = 1 EUR**.

The decision screen is shown on the next page. It has two areas: a test area and a decision area. The upper half is the test area. Here you can test and see the relationship between selected numbers and costs within an interaction and between the three interactions. Use the slider to select a number. Below each slider, the costs for the selected number will be displayed. The graph illustrates the relationship between the selected number and the corresponding costs for every possible number in an interaction.

The bottom half is the decision area. Here you will enter your decisions, which will be relevant for your earnings. By pressing the button "Calculation of Costs", you will be able to see the costs that correspond to your selected numbers. You will also be reminded about the amount of money the participant with the higher and the lower result will receive. After you select a number for each interaction, you will press the button "Next".

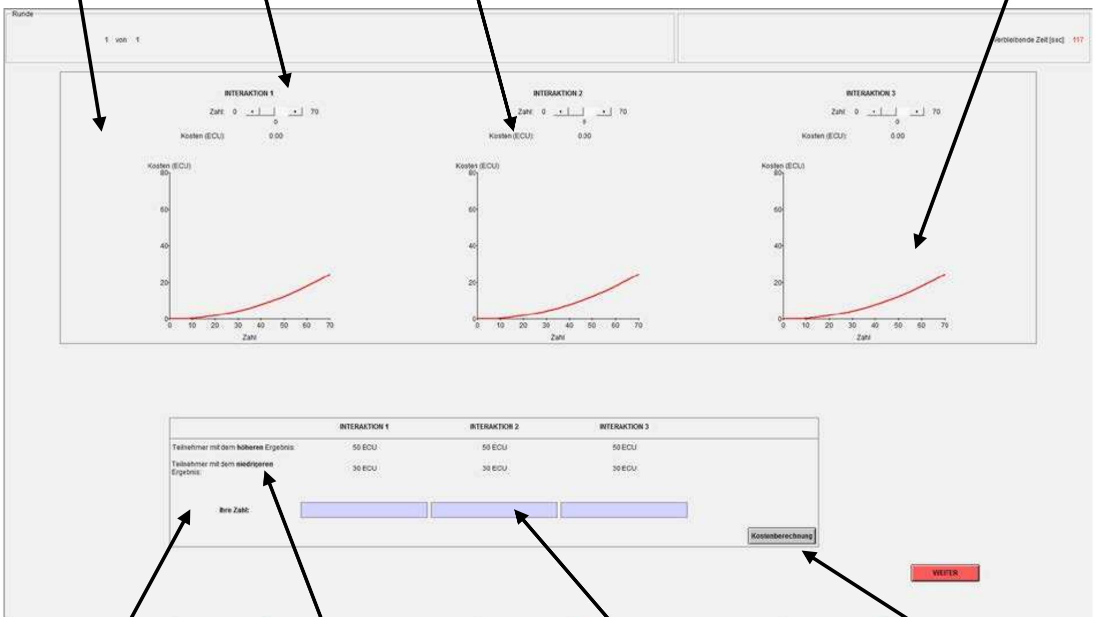
If you have any questions now or during the experiment, please raise your hand and an experimenter will come to you.

Test area

Slider

Under each slider, the costs for the selected number are displayed.

The graph illustrates the relationship between the selected number and the corresponding costs for every possible number.



Decision area

Information about the amount of money the participant with the higher/ lower result will receive.

Enter your number for interaction 1, 2, and 3 in the empty boxes.

Press the button "Calculation of costs" to have calculated the costs for your selected numbers.

References

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental economics 10, 171–178. doi:10.1007/s10683-006-9159-4.

Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association 1, 114–125. doi:10.1007/s40881-015-0004-4.

Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. American economic review 92, 1644–1655. doi:10.1257/000282802762024700.