NBER WORKING PAPER SERIES

ON THE ROLE OF GROUP SIZE IN TOURNAMENTS: THEORY AND EVIDENCE FROM LAB AND FIELD EXPERIMENTS

John A. List
Daan Van Soest
Jan Stoop
Haiwen Zhou

Working Paper 20008
http://www.nber.org/papers/w20008

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2014

Thanks to Steve Levitt for excellent comments on an earlier version of this paper, and also to seminar participants at Harvard University, Princeton University, McGill University, Universidad Autonoma de Madrid, the University of Exeter, Ecole Polytechnique, Paris and the University of California-Santa Barbara for helpful comments. Jonathan Alevy, Roel Ikink, Stef van Kessel, Menusch Khadjavi and Michael Price provided excellent research assistance. Furthermore, we would like to thank NWO and the Erasmus Trust Fund for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

On the Role of Group Size in Tournaments: Theory and Evidence from Lab and Field Experiments
John A. List, Daan Van Soest, Jan Stoop, and Haiwen Zhou
NBER Working Paper No. 20008
March 2014
JEL No. C9,C91,C93,D47,H41

ABSTRACT

Both private and public organizations constantly grapple with incentive schemes to induce maximum effort from agents. We begin with a theoretical exploration of optimal contest design, focusing on the number of competitors. Our theory reveals a critical link between the distribution of luck and the number of contestants. We find that if there is considerable (little) mass on good draws, equilibrium effort is an increasing (decreasing) function of the number of contestants. Our first test of the theory implements a laboratory experiment, where important features of the theory can be exogenously imposed. We complement our lab experiment with a field experiment, where we rely on biological models complemented by economic models to inform us of the relevant theoretical predictions. In both cases we find that the theory has a fair amount of explanatory power, allowing a deeper understanding of how to effectively design tournaments. From a methodological perspective, our study showcases the benefits of combining data from both lab and field experiments to deepen our understanding of the economic science.

John A. List
Department of Economics
University of Chicago
1126 East 59th
Chicago, IL 60637
and NBER
jlist@uchicago.edu

Jan Stoop
Department of Applied Economics
Erasmus School of Economics
PO Box 1738
3000 DR Rotterdam
The Netherlands
stoop@ese.eur.nl

Daan Van Soest
Tilburg University
d.p.vansoest@uvt.nl

Haiwen Zhou
Department of Economics
Old Dominion University
hzhou@odu.edu

Luck is the residue of design—John Milton, 1628[1]

1. Introduction

Tournaments are ubiquitous. Grades in school, receiving a job offer, dating, winning business contracts, and just about everything in between has an element of relative assessment. For instance, internal job promotions in the workplace are decided by relative assessment of the performances of all candidates – just as the relative speed of athletes determines the winner in a running event (Lazear and Rosen 1981). And, patent races can be viewed through the tournament lens as well, where the firm whose invention gets patented first receives most (if not all) of the benefits, while its competitors receive little or no return (Che and Gale 2003).

Over the past several decades, important theoretical work has clarified the advantages and disadvantages of tournaments in which competitors are rewarded according to relative performance (for early work, see e.g., Lazear and Rosen (1981), Holmstrom (1982), Carmichael (1983), Green and Stokey (1983), Nalebluff and Stiglitz (1983), Malcomson (1984) and O'Keefe et al. (1984)).[2] Empirically testing the theoretical predictions of contest (or tournament) models has taken two quite distinct paths: regression-based methods that focus on outputs, and laboratory experiments that are able to measure inputs. One clever illustration of the former method is due to Ehrenberg and Bognanno (1990), who find strong support consistent with the notion that golfers on the PGA tour respond to the level and structure of prizes in tournaments. The seminal laboratory experiment is due to Bull et al. (1987), who find that effort levels converged to theoretical predictions in aggregate, but that individual effort level choices were quite noisily distributed around the equilibrium prediction.

[1] Many scholars credit English poet John Milton (1608-1674) for this quote—specifically "At a Vacation Exercise in the College" (1628)—but the saying does not seem to appear in any of Milton's writings. Branch Rickey, a Major League Baseball Executive is also credited with making this statement in 1915.

[2] The literature has provided several reasons for employing contests, including reducing monitoring costs, dealing with indivisible rewards, and minimizing risks from common uncertainties, but contests do entail potential dangers. For example, they may elicit an incorrect level of effort (moral hazard) or induce the wrong agents to participate (adverse selection); see Lazear and Rosen (1981) and O'Keefe et al. (1984).

To date, the literature has identified the prize structure and the accuracy of the monitoring system – the stochastic element – as the key means to provide the correct incentives in contests. In this study, we enhance the principal's choice set by showing that the number of contestants allowed to compete also has important effects on individual effort levels. On the one hand, larger tournaments induce greater levels of competition and hence we can expect greater effort levels. On the other hand, the probability of winning a large tournament is smaller, providing agents with weaker incentives to exert effort. Loury (1979) and Nalebuff and Stiglitz (1983) argue that the former effect tends to dominate the latter (unless it becomes optimal for some contestants to drop out of the tournament altogether and just collect the loser's prize). Prendergast (1999, p. 35) comes to the opposite conclusion, stating that the size of the prize needs to increase with group size to prevent effort from falling. Complementing the theory, the experimental laboratory results are ambiguous. While Harbring and Irlenbusch (2003) observe effort to decrease with group size, Orrison et al. (1997) find that effort level does not change.

We start from the observation that the role of luck (the random shock) in choosing effort is under-researched in the tournaments literature, and propose that its distribution crucially affects whether effort increases with group size. The order of moves in tournament settings is that agents decide what amount of effort they wish to invest, and then each agent is informed about his 'luck' – where every agent's 'luck' is drawn from the same distribution. An agent's performance is the result of both effort and luck, and the winner of the tournament is the individual with the best performance.

The standard assumption is that each agent's idiosyncratic shock is drawn from a uniform distribution – yet, in many settings the distribution of this stochastic component might be skewed. Consider the case of patent races. Typically there are multiple ways to develop a new product. Suppose that there are many avenues that look very promising (and equally so) ex ante, and that the number of firms participating in the race is large. All firms arbitrarily choose a development path from the available set. Each individual firm then expects to be successful in developing the new product itself, but it also expects that at least one other firm is likely to be successful too. Hence, effort is crucial in determining the winner, and firms invest a lot – and more so than if the number of competing firms is few.

Alternatively, suppose that firms think that the chances of developing a very successful product are small. In that case, they expect that luck will play a crucial role in selecting the winner. If there are infinitely many firms, each firm knows that at least one firm will be so lucky to choose the best approach and win the patent, but it also knows that it is unlikely that it will be that lucky firm itself. Hence, luck is much more important in selecting the winner than effort, and hence firms invest little. The smaller the number of contestants, however, the smaller the probability that one (or more) of the contestants have chosen the best development path, and hence effort may still select the winner – but only if the number of contestants is sufficiently small.

As a real world example, consider the development of medicinal drugs by means of synthesizing chemical components obtained from plants (see for example Simpson et al. 1996). If firms know that a specific substance is likely to be successful in curing a disease, they can target specific plant genera or even species and, ex ante, the odds of developing a new species are large for all innovators. If, on the contrary, little is known about the structure of the successful component, sampling of plants is more random, and the odds of a pharmaceutical company developing the cure are, ex ante, quite small.

Our theory highlights that if the distribution of the uncertainty component is skewed, the number of competitors allowed in the competition has a critical influence on equilibrium effort levels. In particular, as the size of the tournament expands, a contestant's equilibrium effort level (i) decreases if there is little mass on good 'luck', (ii) remains the same with a uniform density, and (iii) increases if there is considerable probability mass on good outcomes. The intuition is that the marginal benefit from committing effort critically depends on both the number of competitors and on the mass associated with "good draws," which depends on the slope of the density function.

Our first test of the theory utilizes a laboratory experiment. By studying experimental markets that differ only in the shape of the distribution of the idiosyncratic shock component, we are permitted a unique insight into whether our theoretical predictions are found in a controlled environment. Experimental methods in the lab thus allow us to study effects that would be quite difficult to identify in naturally-occurring data. Consistent with our theoretical predictions, we find effort does not fall with group size in the case of an increasing density function (implying that the mass on good outcomes is large), whereas it does fall for the decreasing density function (with little mass on good outcomes). When controlling for subjects' risk preferences and for the temporal pattern, we even find that effort *increases* in group size with an increasing density function – providing strong support for our theory.

Our second empirical investigation—a field experiment—continues to rely on randomization to test our theory, but proceeds in a slightly different spirit. Whereas in our laboratory experiment we impose all of the underlying assumptions of the theory and explore effort choices, in the field we test whether these results continue to obtain where simplifying assumptions are not guaranteed to hold. Importantly, however, to test our theory it is necessary to have an observable measure of individual effort and a firm grasp of the underlying distribution of shocks.

While finding environments where these criteria are met is difficult, our search for an appropriate environment concluded when we obtained an agreement with a Dutch commercially-run recreational fishing outfit.[3] Agents in this environment commonly compete in tournaments, and we can measure their effort levels in a straightforward, natural, manner. As such, our simple experimental manipulations are viewed as normal by participants, and with the spatial arrangement of competitors around the lake we have priors on the shape of the uncertainty component's distribution. Fishermen fish for Rainbow Trout, a species that biology has taught us typically school (Liao et al. 2003), and hence the odds of catching fish critically depends on where the schools are located in the pond. Overall, we find evidence consonant with our theoretical predictions in such an environment—as the number of competitors increases, individual effort levels decline.

We view our results as having import in several domains. First, our theory provides intuition as to how effort is related to group size, and provides a direction into interesting positive and normative implications heretofore not discussed. In doing so, we not only add a tool to enhance mechanism design, but provide insights into current policy debates. For example, if effort is not necessarily decreasing in the

[3] We are grateful to Ad and Thea van Oirschot of "De Biestse Oevers", Biest-Houtakker, The Netherlands, for allowing us to use their ponds for experimentation.

number of contestants, then merger and acquisitions cannot be justified by the argument that concentration is necessary to give firms incentives to conduct research and development.

Methodologically, contrary to studies using naturally occurring data, both our lab and field experiments permit a glimpse of individual effort levels (and not just of output measures; see for example O'Reilly et al. (1988), Ehrenberg and Bognanno (1990), Becker and Huselid (1992), Main et al. (1993), Orszag (1994), Lynch and Zax (1998, 2000), and Eriksson (1999), Boudreau et al. (2011, 2013)). More generally, our study showcases the benefits of combining lab and field experimental data to test economic theory.

The remainder of our paper is organized as follows. Section 2 provides our theoretical model and experimental design. Section 3 discusses our empirical results. Section 4 concludes.

2. Tournament Theory

The theoretical literature on tournaments represents a rich assortment of work with several interesting implications. Much of the literature in the area of labor economics can be traced to the work of Lazear and Rosen (1981), who originally clarified the problem of incentives when competitors are paid on a relative basis. Their theory lends structure to several real-world phenomena, including salary structure in corporations and payouts in sporting events.

Green and Stokey (1983) pushed the argument in an important direction by demonstrating that the optimally-designed tournament dominates other reward systems when a sufficiently diffuse common shock exists. Malcomson (1984) later highlighted certain properties of tournaments by examining incentives in an asymmetric information environment. Proceeding in a somewhat different dimension, O'Keefe et al. (1984) take the structure as given and model the problem as one of intensive and extensive optimality: eliciting the correct level of effort and inducing the correct people to participate. In a labor setting, they argue that with the proper use of monitoring probability and prize structure the moral hazard and adverse selection problems can be solved.

A. Theoretical model: risk neutrality

We assume there are $n \geq 2$ risk neutral contestants exerting effort to produce output. In the typical triangular prize format, the agent who produces the highest level of output wins the contest and receives a reward of W_1 while each of the remaining agents receives a payoff of $W_2 < W_1$. Let μ_i denote a representative agent i's effort level, and μ_k denote the effort of her *k*th rival, $i,k \in \{1,2,...n\}$. Let ε_i and ε_k denote identically and independently distributed random variables which have a distribution function denoted by *F*. Without loss of generality, the support of this function is assumed to be symmetric around zero: $\varepsilon_i \in [-a,a]$. The distribution function is assumed to be continuous and twice differentiable and the corresponding density function is *f*. The realized output q_i of contestant i is defined as

$$q_i = \mu_i + \varepsilon_i . \tag{1}$$

Under these conditions, for agent *i* to win the tournament it is necessary that

$$\mu_i - \mu_k + \varepsilon_i > \varepsilon_k \text{ for all } k \neq i .$$

Assuming symmetry, all rivals' effort is the same, denoted by μ. Given the effort level of her rivals, contestant i's probability of producing the best output is $F^{n-1}(\mu_i - \mu + \varepsilon_i)$ for a given ε_i. Integrating over all possible realizations of ε_i, contestant i's expected probability of winning the contest is $\int_{-a}^{a} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i$. Let $C(\mu_i)$ denote contestant i's cost of effort level μ_i: we assume $C' > 0$ and $C'' > 0$. Thus her expected payoff is

$$W_1 \int_{-a}^{a} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i + W_2 \left[1 - \int_{-a}^{a} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i \right] - C(\mu_i) .$$

Contestant i chooses μ_i to maximize the expected payoff. Assuming an interior solution, the first order condition for contestant i's profit maximization is

$$(W_1 - W_2) \int_{-a}^{a} (n-1) f(\mu_i - \mu + \varepsilon_i) F^{n-2}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i - C'(\mu_i) = 0 .$$

In a symmetric equilibrium, $\mu_i = \mu$ for all *i*, and the above equation reduces to

$$(W_1 - W_2)\int_{-a}^{a}(n-1)f^2(\varepsilon_i)F^{n-2}(\varepsilon_i)d\varepsilon_i - C'(\mu) = 0\,. \tag{2}$$

Using integration by parts, we find that

$$\int_{-a}^{a}(n-1)f^2F^{n-2}d\varepsilon = f(a) - \int_{-a}^{a}F^{n-1}f\,'d\varepsilon\,,$$

and hence

$$d\int_{-a}^{a}(n-1)f^2F^{n-2}d\varepsilon\,/\,dn = \int_{-a}^{a}(-\ln F)F^{n-1}f\,'d\varepsilon\,. \tag{3}$$

From (3), when W_1 and W_2 are fixed, the sign of $d\mu/dn$ is the same as the sign of $f\,'$. Thus, a first proposition follows:

Proposition 1: *The form of uncertainty characterizing the tournament affects the relationship between the number of contestants and equilibrium effort levels. When contestants are risk neutral*, a *contestant's effort decreases (if $f\,' < 0$), remains the same (if $f\,' = 0$), or increases (if $f\,' > 0$) as the number of contestants increases.*

The intuition underlying this result is as follows. When an agent chooses her effort level, she naturally compares the marginal benefits and marginal costs of effort. When the number of contestants increases, the probability of one or more other contestants receiving a very good draw is increasing, and this holds independent of whether the density function is increasing, decreasing, or uniform. The increase in this probability has two competing effects influencing the marginal benefit function. The first effect is that "pure luck" (a good realization of a contestant's random variable) is less likely to determine the winner. The larger the number of contestants, the more likely it is that at least some agents end up with high realizations (for a given distribution), and hence the more important effort is in determining the winner. The second effect is that each individual contestant's probability of having the best luck decreases. With convex effort costs, for a given prize the net marginal benefit of effort increases only if the first effect dominates the second.

Three natural examples are intuitively plausible. First, suppose the density function is increasing on its support. The contestant knows that she has a high probability of receiving a good draw, but she also knows that the probability of one or more other contestants receiving a good draw is increasing in group size. Hence, the larger the group, the closer the contestants are in terms of likely outcomes—good draws. As a result, the first effect dominates the second when the number of contestants increases, and effort plays an important role in selecting the winner.

Second, suppose that at the right-hand side of the distribution the density function is decreasing – as is the case with a normal distribution. The contestant knows that her probability of receiving a good draw is small, whereas the probability of at least one other contestant receiving a good draw increases in group size. Hence, the larger the number of contestants, the smaller the likelihood that putting in extra effort will pay off, and hence the second effect dominates the first. Finally, as is typically assumed in the literature, suppose the density function has a zero slope. In that case the first effect exactly cancels the second effect and effort does not change with the number of contestants. Hence, the form of uncertainty characterizing the tournament determines whether equilibrium efforts increase, decrease or stay the same if the number of contestants in the tournament increases.

A simple numerical example facilitates interpreting the results. Consider the case where there are only two possible outcomes; one can have either a good draw (with probability p) or a bad draw (with probability $1-p$). Note that for $p >$ ($<$) 0.5, the mass on good outcomes is larger (smaller) – as is the case with $f' > 0$ ($f' < 0$). Furthermore, assume that the costs of effort and the size of the prize are such that it is not profitable for a contestant to exert any effort if she herself receives a bad draw and at least one other contestant receives a good draw. Therefore, effort only plays a role in selecting the winner if (i) either the contestant receives a good draw *and* at least one other contestant also receives a good draw (which happens with probability $p(1 - (1-p)^{n-1})$), or (ii) the contestant receives a bad draw *and* none of the other contestants receives a good draw either (the probability of which equals $(1-p)^n$). So the probability that effort matters is $\pi(n,p) = (1-p)^n + p(1 - (1-p)^{n-1})$. If $n = 2$, $\pi(n,p) = (1-p)^2 + p^2$, which is a U-shaped function in p, with a minimum at $p = 0.5$. So, the more (or less) mass there is on good outcomes, the more likely it is that the contestants are close in terms of 'luck', and hence the marginal benefits of investing effort are larger. If $n \to \infty$, $\pi(n,p) = p$, and the odds that effort matters increases linearly in the amount of mass on good outcomes. With $n \to \infty$, the probability that at least one other contestant receives a good draw is equal to 1. Hence, the higher the probability that the decision maker receives a good draw too, the higher the probability that effort will be decisive in determining the winner. So, $\pi(2,p) > (<) \pi(\infty,p)$ if $(1-p)^2 + p^2 > (<) p$, and hence the probability that effort matters decreases (increases) in group size if $p < (>) 0.5$. Also note that for $p = 0.5$ we have $\pi(2,p) = \pi(\infty,p)$, implying that in this instance the marginal benefits of effort remain constant if group size changes.

B. Theoretical model: risk aversion

One key assumption in the previous subsection is that agents are risk neutral – which follows the literature closely. In this sense, any empirical test represents a joint hypothesis test—risk neutrality and equilibrium play. In an effort to extend this aspect of the literature, we consider our theoretical predictions when agents are risk averse. This appears to be a natural assumption, as recent explorations of individual risk preferences (e.g., Holt and Laury, 2002) suggest that a majority of agents act in a manner consistent with a model of risk-aversion when confronted with choices of lottery payoffs that are typical in lab experiments.

Continuing with our theoretical model described above, we denote the well-behaved utility function for an agent as U: $U' > 0$, $U'' < 0$. An agent's expected utility is therefore:

$$U(W_1 - C(\mu_i))\int_{-a}^{a} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i$$

$$+U(W_2 - C(\mu_i))\left[1 - \int_{-a}^{a} F^{n-1}(\mu_i - \mu + \varepsilon_i) f(\varepsilon_i) d\varepsilon_i\right].$$

In a symmetric equilibrium, the condition for profit maximization becomes

$$V \equiv ((U(W_1 - C) - U(W_2 - C))\int_{-a}^{a}(n-1)f^2(\varepsilon_i)F^{n-2}(\varepsilon_i)d\varepsilon_i$$

$$-\frac{1}{n}\left[U'(W_1 - C) + (n-1)U'(W_2 - C)\right]C'(\mu_i) = 0\,. \tag{4}$$

From (2) and (4), a contestant will commit less effort when she is risk averse if the following inequality holds:

$$W_1 - W_2 > \frac{n(U(W_1 - C) - U(W_2 - C))}{U'(W_1 - C) + (n-1)U'(W_2 - C)}\,. \tag{5}$$

Since the utility function is concave, (5) always holds. Thus, ceteris paribus, risk averse agents will commit less effort than risk neutral agents. The intuition behind this result is as follows. Given that the utility function is concave, $U(W_1) - U(W_2)$ is smaller than $W_1 - W_2$; thus, as a risk averse agent's valuation of the benefits of winning is smaller than of a risk neutral agent, the effort level chosen by the former is smaller.

From (4), the relationship between a contestant's effort and the number of contestants is given by

$$\frac{d\mu}{dn} = -\frac{\partial V / \partial n}{\partial V / \partial \mu}\,.$$

From the second order condition for a contestant's payoff maximization, $\partial V / \partial \mu$ is always negative. Thus, $d\mu / dn$ has the same sign as $\partial V / \partial n$. Partial differentiation of (4) yields

$$\frac{\partial V}{\partial n} = [U(W_1 - C) - U(W_2 - C)]\int_{-a}^{a} f' F^{n-1}(-\ln F)d\varepsilon \; - \frac{1}{n^2}\left[U'(W_2 - C) - U'(W_1 - C)\right]C'\,. \tag{6}$$

With diminishing marginal utility the second term on the RHS of (6) is always negative, while the first term on the RHS may be negative, zero or positive depending on whether the density function is decreasing, uniform, or increasing. For the increasing density case, $f' > 0$, a risk-averse agent's equilibrium effort level is ambiguous over changes in the number of rivals. As noted above, for $f' > 0$, an agent's effort level increases with the number of contestants when they are risk neutral. With the additional effect from risk aversion, the relationship becomes ambiguous because risk aversion decreases the utility payoff from winning. In much the same manner, for $f'=0$, a risk neutral contestant's equilibrium effort level does not change as the number of contestants changes, but introducing risk aversion causes this relationship to become negative. This gives rise to our second proposition:

Proposition 2: *Under risk aversion*, a *contestant's effort decreases with group size if the density function is decreasing or uniform, whereas the effect of group size is ambiguous for increasing density functions*.

3. Experimental Evidence

To test our theory, we proceed in two complementary directions. We begin by imposing the major assumptions of our theory in a laboratory experiment, allowing an analysis of the effects of alternative shock distributions on individual effort levels. We proceed to an environment—a field experiment—where we can only be certain that a few of the critical assumptions are met: those that provide enough structure to provide theoretical predictions. The proceeding discussion will be organized as an explanation of each experimental method followed by the empirical evidence drawn from that approach.

A. Laboratory Experiment

In our lab experiment subjects compete in tournaments with either 1 or 3 other contestants (and hence in tournaments with group size 2 or 4). They choose effort μ (or a "decision number" in the language used in the instructions) by selecting a number between 0 and 100. Subjects are presented with costs associated with each decision number, using the following cost function: $C(\mu)=\frac{3}{10{,}000}\mu^2$. Costs are measured in points, and hence the costs of picking a number between 0 and 100 range from 0 to 3 points.

Output (referred to as the "total number" in the instructions) is the sum of the decision number chosen by subject *i*, μ_i, and a "random number", ε_i, randomly drawn from a distribution with support [-100,100]. The subject with the highest output $\mu_i + \varepsilon_i$ in his/her group in a tournament is the winner. Each subject *i* is informed about the realization of her draw, ε_i, only after having chosen μ_i. Random numbers are drawn from decreasing, uniform and increasing density functions that are specified as $f(\varepsilon) = -\frac{\varepsilon}{20{,}000} + \frac{1}{200}$, $f(\varepsilon) = \frac{1}{200}$, and $f(\varepsilon) = +\frac{\varepsilon}{20{,}000} + \frac{1}{200}$, respectively.[4] To ensure the second order condition for a contestant's payoff maximization is satisfied regardless of whether there are two or four contestants, we set $W_1 = 5$ and $W_2 = 2$ points. A point is worth one euro.

We implement tournaments of $n = 2$ and $n = 4$ for each of the three density functions. We denote the two-person decreasing density function tournament treatment as D2; other treatments use similar acronyms. We ran six sessions with a total of 136 subjects recruited from Tilburg University's student body. In every session, subjects participated in two series of ten tournaments. Within a session the density function was held constant, but the two series of tournaments were implemented with group sizes of 2 or 4; the order in which they were played was balanced across sessions. For example, one session implemented D2 first and then D4, while another session first implemented D4 and then D2 – and similarly for the uniform and increasing density functions. Acronyms for these sessions are D2D4 and D4D2, respectively. This design allows us to not only compare effort levels across treatments (in a between-subject analysis), but also to analyze individual changes in effort induced by changes in group size (in a within-subject analysis). Within a series of ten tournaments group composition was kept fixed, but groups were randomly rematched after the first series had been completed. The experiment was compiled and conducted with z-Tree (Fischbacher 2007) .

---

4 The associated distribution functions are $F(\varepsilon) = -\frac{\varepsilon^2}{40{,}000} + \frac{\varepsilon}{200} + \frac{3}{4}$, $F(\varepsilon) = \frac{\varepsilon}{200} + \frac{1}{2}$, and $F(\varepsilon) = \frac{\varepsilon^2}{40{,}000} + \frac{\varepsilon}{200} + \frac{1}{4}$.

In the sessions, subjects were informed of the number of competitors, the structure of payoffs and costs, and the structure of the density function from which shocks were drawn (the instructions are in Appendix II, to be made available online). They were also informed that they would participate in a series of ten tournaments of a specific type, but that after the tenth tournament was completed one of the ten would randomly be selected to determine their payouts. The same procedure was followed in the second tournament type that they participated (i.e., with the same density function, but with larger or smaller group size). Within each of the ten tournaments in a series, decision numbers were chosen simultaneously, random numbers were independently drawn, and subjects were informed of all effort levels chosen and random numbers received by all subjects in their group – and also whether they had the highest output in a round, or not.

Our theory importantly highlights the effect of risk preferences—in particular, the assumption of risk neutrality versus risk aversion plays an important role. At the end of every session we elicited our subjects' risk attitudes using the method developed by Holt and Laury (2002). (The instructions are included in Appendix III, to be made available online). This approach follows Lange et al. (2007). As Lange et al. (2007) note, attempting to measure risk postures in one game and applying them to more closely explore behavior in another is not novel to this study (see, e.g., Eckel and Wilson, 2004). Yet, there are important issues with such an approach. First, whether risk preferences are stable across games, over time, etc., is an open question. Second, whether individual unobservables that influence risk posture are correlated with behavior in the tournament experiment is unknown. For these and other reasons, because risk posture is not assigned randomly across players (such as the number of competitors or the luck component in the tournament) an important caveat must be placed on the results from such an exercise. Thus, when interpreting the empirical results from this part of the experiment these factors should be considered.

To put these data to use, we first calculated the average coefficient of risk aversion of all participants. On average, they made almost 6 safe choices, yielding a coefficient of Constant Relative Risk Aversion (CRRA) of about 0.5. Next, using (2) and (4) respectively, we calculate the equilibrium effort levels assuming (i) that all contestants in a tournament are risk neutral, and (ii) that all contestants in a group are endowed with a CRRA equal to 0.5.[5] The theoretical point predictions of equilibrium effort are presented in Table 1. Rows in Table 1 represent whether the treatment was carried out with 2 or 4 competitors in the contest, and columns denote the shape of the density function from which the idiosyncratic shocks are drawn: decreasing $(f'<0)$, uniform $(f'=0)$, or increasing $(f'>0)$.

<Insert Table 1 here>

Consistent with the results in sections 2A (including the simple binary example therein) and 2B, Table 1 indicates that for n = 2, the equilibrium effort level is lowest in case of the uniform density function, while for n = 4 the effort level with a uniform distribution is larger (smaller) than that with a decreasing (increasing) density function – independent of whether agents are assumed to be risk neutral, or risk averse. So, the predicted patterns are the same when assuming risk neutrality or risk aversion: independent of risk attitudes, equilibrium effort is U-shaped when moving across the three distribution functions from $f' < 0$ via $f' = 0$ to $f' > 0$ in case of $n = 2$, while the relationship is strictly increasing for n = 4 in case agents. Yet risk attitudes do have important consequences for the magnitude of the predicted change in effort when changing group size. Independent of risk posture, the predicted fall in effort is substantial in case of a decreasing density function, but the predicted increase with the increasing density function is much smaller when agents are assumed to be risk averse than if they are assumed to be risk neutral. And while risk-neutral agents are not expected to change their effort levels in the uniform density case, risk-averse agents are expected to reduce slightly their effort level when group size increases.

Having derived these predictions, the question is how groups with heterogeneous risk preferences are expected to play. Intuitively, the pattern of effort levels should be closer to the predicted equilibrium levels assuming risk aversion than to those assuming risk neutrality. First, eighty percent of our subjects

[5] To calculate Nash effort levels under risk aversion, we assume that $U(x) = x^{1-r}/(1-r)$, where $r = 0.5$.

are measured to be risk averse, and second, because risk averse players put in less effort than the risk neutral ones, the risk neutral players will put in less effort than the equilibrium prediction for their type. If risk-neutral subjects play against other contestants putting in relatively little effort, best response is to put in less effort too (Bull et al. 1987). On the basis of this and of Table 1, we postulate three hypotheses.

Hypothesis 1: *In the two-player tournaments, the observed effort levels for the decreasing and increasing density functions should be similar, and higher than the observed effort level when the density function is uniform.*

Hypothesis 2: *Comparing the effort levels for n=2 and n=4 and because risk-averse subjects are expected to be present in the subject pool, we expect effort to fall (significantly) in case of decreasing and uniform density functions (with the decrease in the former to be larger than in the latter). In case of an increasing density function, effort does not decrease (significantly) when increasing the number of contestants, and may even increase.*

Note that hypothesis 1 is expected to hold if the random allocation of subjects to groups is such that group composition is similar. Hypothesis 2 is more cumbersome because the size of the decrease in case of $f' = 0$ and the change in effort in case of $f' > 0$ crucially depends on the number of risk neutral and risk averse subjects in the various groups.

Experimental Results of the Lab Experiments

Table 2 presents a summary of the experimental results. To aid in the analysis, we have pooled the observations from the same treatment in different sessions, averaging effort over all subjects and all periods (or tournaments). For example, the D2 data are obtained from the first ten tournaments of D2D4 and of the last ten tournaments of D4D2.

<Insert Table 2 here>

The treatment effects presented in Table 2 are in line with the theory, although not all differences in effort are statistically significant. And, the levels themselves are quite close to those predicted in Table 1 (although the mean effort level in D2, μ_{D2}, is considerably larger than predicted, while μ_{I4} is lower). In line with Hypothesis 1, μ_{D2} and μ_{I2} are larger than μ_{U2} – mean effort is indeed higher if the contestants in a two-person tournament are likely to be close in terms of luck. Contrary to the theoretical predictions, however, we also find that $\mu_{D2} > \mu_{I2}$. This is surprising because D2 and I2 are mathematically equivalent – the difference between μ_{D2} and μ_{I2} is thus due to just a framing effect.

Next, as predicted by Hypothesis 2, mean effort is higher in I4 than in U4, and it is also higher in U4 than in D4. In I4, at least one other subject is expected to get a good draw, but each subject expects to receive a good draw for him/herself too – and hence effort is expected to be important in selecting the winner. In D4, effort is less likely to be decisive because here the probability of receiving a good draw are small, while the chances that at least one of the three others receives a good draw, is still quite large.[6]

Overall, the empirical results support the theory quite well, and the average effort levels are even quite close to the point predictions presented in Table 1. However, our main interest is in the effect of group size on effort levels. In line with the theory, we find that mean effort is significantly larger in D2 than in D4, with $p < 0.01$ (according to a standard Mann-Whitney U test). Next, while we find that μ_{U4} is 2.5 units below μ_{U2}, the difference is not statistically different, and this holds even more for the (negligible) difference between μ_{I2} and μ_{I4}. The latter two results are in line with the presence of risk-averse individuals in the subject pool.

The results thus provide considerable support for the theory. Yet the averages may hide important temporal patterns, and one may wonder how risk attitudes affect the results, and whether learning takes place over the iterations. To test for these features, we estimate a simple OLS model:

[6] In terms of statistical significance, the relevant Mann-Whitney U tests yield the following. With respect to the two-person tournaments, we find $\mu_{D2} > \mu_{U2}$ at p =0.018, and also $\mu_{D2} > \mu_{I2}$, albeit at p = 0.080 only. Regarding the four-person tournaments, we have $\mu_{I4} > \mu_{D4}$ at p =0.073, but similar tests allow us to neither reject that $\mu_{D4} = \mu_{U4}$, nor that $\mu_{U4} = \mu_{I4}$.

$$\mu_{ijt} = \alpha + \sum \beta_j \, T_j + \gamma \, RA_i + \delta \, t + \eta \, PlayedFirst_{ij} + \varepsilon_{ijt}, \quad (7)$$

where μ_{ijt} is the effort level chosen by subject *i* in tournament *t* of treatment *j*, T_j are treatment dummies, RA_i is subject *i*'s switching point in the Holt-Laury test implemented at the end of the session, and $PlayedFirst_{ij}$ is a dummy variable that takes the value 1 if treatment *j* is the first treatment implemented in the session subject *i* participates in, and zero if it is the second. We run the regressions using all observations, but also when selecting only those in period 5, or higher. The regression results of these OLS models are presented in Table 3.[7] In the table, the intercept captures the mean effort level in U2.

<Insert Table 3 about here>

Focusing on the regression results using all observations as presented in column (i) of Table 3, we find that effort tends to be higher in the treatment that is played first in a session (as shown by the coefficient on "Treatment played first"), risk-averse individuals tend to choose lower effort levels than risk-neutral individuals, and effort does not tend to decrease appreciatively over the periods.

Controlling for these factors, the estimated coefficients on the treatment dummies show a pattern consistent with the theory. Both the coefficients on D2 and I2 are significantly different from zero, implying that effort is higher in these two treatments than in U2. In addition, effort in D4 is significantly smaller than in U2, while effort in U4 is not significantly smaller than that in U2. Finally, we also find that effort in I4 is higher than in U2 (albeit at $p < 0.10$ only) – and hence it is a fortiori higher than U4 too. When just focusing on the observations for period 5 and higher, as shown in column (ii) of Table 3, the treatment differences tend to increase.

B. *Field Experiment*

[7] Our results are unaffected when running Tobit models with censoring at effort levels 0 and 100.

The lab data thus paint a picture that is consonant with our theory. As a next test of the model's predictions, we take our theory to the field where many of the theoretical assumptions cannot be assured to be met. Over the past two decades, a rich assortment of tournament studies in field settings have arisen, shedding important insights on relevant economic models. These studies revolve exploiting naturally-occurring data, and exclusively deal with variables concerning outputs rather than inputs. For example, in sport settings, Ehrenberg and Bognanno (1990) and Orszag (1994) report golf scores, Becker and Huselid (1992) report speed and outcomes in auto racing, while Lynch and Zax (1998) report outcomes of a horse race to measure effort in a tournament. Studies of tournaments within firms use a similar approach: O'Reilly et al. (1988) use measures of sales, profits and number of employees to explain CEO wages; Main et al. (1993) extend this line of work in a similar spirit. Eriksson (1999) takes a different approach, but also one that essentially measures outputs, as he assumes effort is equal to average profits divided by sales. Boudreau et al. (2011, 2013) look at naturally-occurring data on the effects of group sizes in tournaments on software development. Their variable of interest is the score assigned to a solution of a software problem.

Rather than focusing on naturally-occurring data, we move this literature in a new direction by making use of a field experiment, wherein we can measure inputs. In doing so, it is important to craft an experimental design that exogenously varies our major treatment variable—number of competitors—in an environment that permits an understanding of the other important features of the situation. This approach provides us with an opportunity to observe behavior of agents who have endogenously selected into the market, while simultaneously making use of controls afforded by an experiment. To this end, we strived to exploit a naturally-occurring environment whereby the random stochastic component takes a shape that is well understood by the participants.

Finding such an environment is not trivial, but our search concluded when the operator of a recreational fishing outfit in The Netherlands agreed to provide i) access to its customers and ii) space on the ponds to carry out our experiment. The outfit consists of three rectangular fishing ponds, each of which is roughly 8500 square feet. The normal procedure is that customers pay an entry fee of 12.50–15

Euros and fish for a period of 4 to 5 hours, depending on the season. The entry fees for these ponds vary as a function of the type and number of stocked fish – rainbow trout or salmon trout. Further, customers do not have 'property rights' regarding the fish that are thrown in on their behalf; rather, at each pond there is no "institutional" constraint on catch. Any fish caught needs to be taken home; throwing them back is prohibited to protect the health of the remaining stock.

This setting has several features that are ideal for our purposes. First, it provides us with a participant pool that naturally competes in tournaments, and indeed some of our subjects have participated in national fishing competitions. Second, the fishing technology is geared towards exploiting the behavior of prey by continuously casting and reeling. Bait is thus dragged through the water, seducing the trout to chase and take. Although theoretically reeling in too quickly means that the trout is outrun, we show below that the number of fish caught per time period is an increasing function of the number of casts in that period. This suggests that we have a measure of effort—the number of casts per period—that is a useful measure to test our theory. Clearly, the same casting frequency may imply very different effort levels for different subjects. Thus, to account for skill heterogeneity our design must be careful to provide within–subject treatment variability.

Third, the fishing pond permits us a natural test of our theory for the case of a decreasing density function. Biological models inform us that trout fish school (Liao et al. 2003). While in principle the pond is small enough that fishermen could cast their bait to where a school is located, the standard regulations of the fishing facility require customers to only fish in the rectangle in front of them to avoid lines getting tangled up – fishermen are not allowed to cast their bait farther away than half the width of the pond (remember that fishing always takes place on the long sides of the pond), and not farther left and right than half the distance to their neighbors. The combination of this rule and the fact that trout school, implies that the density function of 'luck' is decreasing: schools never cover more than just a few rectangles, and hence the amount of mass on having good luck is quite small. Also, luck is essentially independent: a school can be located (at a particular moment) in the rectangle of one contestant in a tournament, it can be located on the shared perimeter of the rectangles of two or more contestants (implying that two or more contestants have good luck) or it can be located in none. Hence, 'luck' is drawn largely independently between fishermen (also note that we always replenish the stock of fish at the beginning of every new tournament in order to reduce the negative externalities associated with catching fish). But because there are always 16 participants at every session, the chances of having really good luck are quite small. Statistical support for (i) the relationship between effort and catch and (ii) the argument that there is little mass on very good outcomes (or, that is negatively skewed), is provided in Appendix I.

With these advantages, of course, come disadvantages. First, the field data are likely to be even noisier than the laboratory data. Because of that reason, we decided to increase the difference in group size, having subjects participate in tournaments of eight rather than of four. Second, there is natural heterogeneity in the population not only in terms of risk attitudes (as was the case in our laboratory experiments), but also in terms of skills and the cost of effort.[8]

Third, whereas standard tournament theory is static, our field setting is dynamic in the same spirit as the empirical studies using naturally-occurring data cited above. This is important because the pond is small enough that all subjects can monitor the number of fish caught by the other participants. If a participant notices that another participant is doing very well, he may decide to increase his effort, or to give up – depending on how many fish he caught himself. If we do not disclose who competes with whom in the same tournament, the probability that one competes with that very successful participant is 1/15 in case of n = 2, while it is 7/15 in case of n = 8. To avoid effort decisions to be incomparable across the two treatments, we decided to inform the subjects with whom they were competing in each tournament.[9]

[8] The reader may argue that recreational fishermen enjoy fishing, and hence that effort is costless. Note, however, that the tournament setting induces fishermen to put in more effort than what they normally do. That means that the utility of the act of fishing itself is a hump-shaped function of effort. Without other incentives fishermen would choose the effort level associated with the utility function's peak, while the costs of putting in more effort is equal to the disutility associated with fishing harder than usual.

[9] We can address the issue of the game being dynamic rather than static (at least to some extent) by not only analyzing average effort over the entire tournament, but also average in just the first 15 minutes – because the

Although these differences are not exhaustive, they highlight that field experiments present a tradeoff: they give up some of the controls of a laboratory experiment (such as induced valuations, or robots guaranteed to play equilibrium strategies against human subjects – cf. Bull et al. 1987) in exchange for increased realism. In this manner, our field experiment matches the real-world settings which tournament theory attempts to explain: our fishermen are not told explicitly the distributions of other's valuations and they have previous experience in this environment. In this manner, the exploration provides a useful middle ground between the tight controls of the laboratory and the vagaries of completely uncontrolled naturally-occurring data.

The execution of the tournament experiment was straightforward and followed four steps. First, sports fishermen were recruited via a registration list during the week previous to the planned session. Second, upon arrival, we explained the experimental instructions in a quiet area removed from the other customers. In the instructions we explained that we rented a specific pond and that each subject will participate in 4 tournaments with an alternating number of other competitors—either 1 or 7 per tournament (the instructions are presented in Appendix IV, to be made available online). Every tournament lasts exactly one hour, and the winner of a tournament is the person who catches the most fish during the hour. The fishermen were told that the winning prize is 10 euro's, independent of group size. In case of a draw, the winner is determined by whoever caught the first fish. In case of a tie, we flip a coin to determine the winner.[10]

Third, shortly before the first tournament we stocked the pond with 58 rainbow trout. Recall that the recreational fishing outfit does not allow throwing back any fish that has been caught. Therefore, to make sure that the number of fish in the pond was always the same at the beginning of each tournament,

---

difference in the number of fish caught between contestants is smaller in the first 15 minutes than at the end of the tournament.

[10] We were careful to follow the rules applied by the Dutch Trout Fishing Championships except for the tie breaker. In the official Championships the total weight of the fish caught determines the winner in case of a tie. This is not feasible in our experiment since each participant is in four tournaments. Breaking ties on the basis of total weight of fish caught in the particular tournament round would imply using four coolers per fishermen to store the fish separately. Each sports fisherman usually just carries one, and hence for practical purposes we used the time elapsed before catching the first fish as a tie breaker.

we threw in the same number of fish as was caught in the previous tournament before the next tournament began. During the stocking process before each new tournament began, we allocated initial fishing spots by means of a lottery. All participants drew a numbered spot tag from a closed linen bag, and moved to their new spots before the next tournament started. Fishing only occurs on the long sides of the pond, and we always use 8 of the 10 fishing spots on each side. A whistle blow marks the beginning and end of each tournament.

The fourth and final step involves participant remuneration. As noted above, the agent received 10 Euro's for each tournament victory, and hence the maximum prize money per subject is 40 Euro's. In addition, each subject received 5 Euro's participation fee. Also, we collected all fish caught and redistributed them lump-sum to session participants.[11] The pecuniary outlay for the experiment, consisting of the costs of fish and the payments to the fishermen, was roughly 600 Euro's per session.[12]

Before discussing our empirical results, a few outstanding issues merit brief mention. First, we ran 4 sessions with 16 participants each session; subjects were allowed to compete in only one session each. Second, given that a within-subject design was necessary, we alternated the group sizes of the tournaments within each session. In sessions 1 and 3 the tournaments were played in group sizes of 2, 8, 2 and 8 in rounds 1-4 respectively, and in sessions 2 and 4 group sizes were 8, 2, 8 and 2. Thus, in sessions 1 and 3 we had eight tournaments of $n = 2$ in rounds 1 and 3 and two tournaments of $n = 8$ in rounds 2 and 4. For session 2 and 4, we have two tournaments of $n = 8$ in rounds 1 and 3, and eight tournaments of $n = 2$ in rounds 2 and 4. Finally, in light of our theoretical model and the stochastic component, under our design we have one comparative static prediction to test: *contestants' effort decreases as the number of*

[11] One might have chosen to not allow participants to take home any fish. We wished to avoid waste (in total 487 fish were caught), and could not give them to a charity due to perishability. We therefore decided to redistribute them equally among all participants in a session. The marginal incentive to catch another fish (apart from the increased likelihood of winning the tournament) is thus 1/16th of its value. Since this marginal incentive is small and independent of treatment we believe it will not affect our treatment estimate.

[12] In sum, the benefits of participating in the experiment are (i) fishing four hours for free, (ii) receiving a show-up fee of 5 Euros, (iii) taking home 1/16 of the number of fish caught in the session, and (iv) earning, in expectation, 12.50 Euros in prize money.

contestants increases. This prediction should hold for both risk neutral and risk averse competitors, hence we do not gather data on individual risk posture.

Experimental Results of the Field Experiment

We now turn to our main interest of measuring the impact of group size on fishing effort. As a first glimpse into the received data patterns, we provide Table 4. Panels A and B in Table 4 include means and standard deviations of the effort levels across treatment categorized by period, without adjusting for the data dependencies – for sessions 1+3 and sessions 2+4, respectively.

<Insert Table 4 here>

The cleanest test of the theory is to compare effort levels in rounds 1 of the sessions 1+3 (in which participants competed in groups of 2) versus those of sessions 2+4 (in which participants competed in groups of 8). The difference in effort levels is striking: the mean level is 0.13 casts per minute higher in the n = 2 tournaments than in the n = 8 tournaments, and this difference is significant at $p < 0.020$ according to the appropriate Mann-Whitney U test. Hence, the first-round data provide between-subject support for the hypothesis that effort tends to be smaller for larger groups if there is relatively little mass on good outcomes.

Next, we can analyze the within-subject response to changes in group size by analyzing how effort changes if group size is increased (see the change in effort in Sessions 1+3 between rounds 1 and 2 as presented in Panel A of Table 4) and if it is decreased (see the change in effort in Sessions 2+4 between rounds 1 and 2; see Panel B). Both effort levels are observed to fall from round 1 to round 2, but only significantly so if the group size is increased (at $p = 0.020$; the p-value associated with the decrease in group size is equal to 0.294). A similar analysis can be undertaken by analyzing the changes between rounds 2 and 3, and again we find that the fall in effort is significant only in case of an increase in group size (as the p-value associated with the change in effort in Sessions 2+4 is equal to 0.088, whereas it is

0.963 in Sessions 1+3). Finally, none of the changes in group size results in a significant change in effort between rounds 3 and 4, independent of whether group size is increased, or decreased.

The within-subject analysis thus suggests that effort tends to fall when group size increases, but that it does not increase when group size decreases. This analysis may not capture the relevant underlying pattern, though, because of two reasons. First, the observed reduction in effort in most rounds may hint at the importance of fatigue. Second, while our theory is static, the analysis presented in Table 4 captures a dynamic process – subjects can observe the success of their competitor(s), which may influence their behavior.

Regarding the pattern of play as shown in Table 4, note that effort levels in sessions 1+3 are invariably higher than in sessions 2+4, and also that they are declining over time – except for the last round. This may suggest that in the first tournament we anchor our fishermen at a certain level of effort intensity (with the intensity being higher in the sessions starting with the $n = 2$ tournaments), that the first increase in group size discourages people from keeping up their effort intensity, but that the impact of fatigue is compensated by the incentives to fish at higher effort levels in case group size decreases.

We explore the temporal pattern explaining the change in individual effort when group size increases or decreases using standard regression analysis. As dependent variables we use the average effort of an individual over the entire one-hour tournament, but also his/her average effort level over the first 15 minutes of every tournament. We do so because behavior in the first 15 minutes of each tournament is expected to be a better fit for the static model presented in section 2.[13] As a tournament progresses, some fishermen may have been more successful in catching fish than their competitors; the

[13] Another way of capturing the static aspect of our theoretical model is by looking at effort levels up to the moment at which the first contestant of a group catches his first fish. Although theoretically appealing, this approach is not suitable given our data. At the start of each new tournament, measured effort is larger than during the rest of the tournament especially because fishermen throw in more frequently and reel their baits in faster in order to be able to adjust their fishing gear to the new spot (e.g. finding the optimal length of line between hook and floater). Effort levels only settle after five or ten minutes, while also quite a few fish tend to be caught in this period. Hence, using effort data up to the first fish caught tends to just reflect the efforts participants make to adjust their fishing gear, and this effort is largely insensitive to the size of the tournament subjects are participating in.

larger the difference in the number of fish caught between the leader and his closest competitor, the lower the amount of effort put in by both (cf. Chan et al. 2009, and see also footnote 9).

Following equation (7), we estimate a panel data model at the individual level. Given the shortness of the panel, we use first differences and analyze changes in effort as follows:

$$dE_{it} / E_{it} = \beta_1 * Decreasegroupsize_{it} + \Sigma\, \phi_s + \varepsilon_{it,} \tag{8}$$

where E_{it} is the effort choice for subject i in period t, dE_{it} is the difference in effort of subject i between period t and period $t-1$, Decreasegroupsize$_{it}$ is a dummy variable with value 1 if group size decreases from 8 in period t-1 to 2 in period t, and 0 otherwise. Furthermore, ϕ_s, s=1,…, 4, are session dummy variables; note that the constant is suppressed. Finally, ε_{it} are standard errors clustered at the subject level. In this model, the session variables $\phi_1, \ldots, \phi_4$ allow for session-specific percentage decreases in effort over time while the dummy *Decreasegroupsize*$_{it}$ captures the common response (if any) in all four sessions to decreases in groups sizes (which occur between periods 2 and 3 in Sessions 1 and 3 and between periods 1 and 2 and between 3 and 4 in Sessions 2 and 4). We run the model using two different dependent variables, the change in effort based on the average amount of effort put in the current and in the previous tournament (see Table 5, column (i)), and the change in effort based on the average amount of effort put in in the first 15 minutes of the current compared to that in the first 15 minutes of the previous tournament (see Table 5, column (ii)).

<Insert Table 5 here>

We find evidence consonant with our theory in both regression models. In both columns, the session dummies reflect that effort tends to fall over rounds during the session – between 9 and 25% in case of the data averaged per one-hour round, and between 15 and 22% in case we just look at how effort in the first 15 minutes of every round. However, the fall in effort is smaller if group size is decreased between rounds as evidenced by the positive coefficient on *Decreasegroupsize* in both regression models (with p = 0.096 in column (i), and p = 0.055 in column (ii)). Hence, if group size is increased, effort unambiguously falls, but the fall is smaller in case group size is decreased – and effort may even increase in some periods.

4. Conclusions

Tournament models have played an important role in the design of organizational reward systems, governmental allocation of resources, sporting events, promotion contests, and innovation contests. Although much progress has been made in understanding the theoretical underpinnings of tournaments, yet the empirical evidence is still scarce to date.

Our study begins by expanding the theoretical literature by exploring how equilibrium effort levels vary with contest design; in particular, the relationship between group size and the idiosyncratic shock component. The theory provides several predictions, perhaps most importantly that nature (the assumed shape of the density function) is critical in determining equilibrium effort levels – when agents are assumed to be risk-neutral, but also if they are risk-averse. In this regard, if the form of uncertainty that characterizes the tournament process is skewed, then equilibrium effort levels depend crucially on the number of competitors. In the case of the standard model in the literature, (uniformly distributed idiosyncratic shocks), the number of competitors is predicted not to influence effort levels (if agents are risk neutral), or only slightly (if agents are risk averse).

We test the theory using complementary lab and field experiments. Our first method is to use a laboratory experiment, which permits us to study markets that differ only in the shape of the density function, allowing a unique insight into whether changes in the component's shape itself can lead to predicted changes in behavior. Lab experimental methods thus allow us to study such effects that would be difficult to identify in naturally occurring data. Our second approach is to maintain randomization, but design an experiment in the field that resembles the important features of our theory and permits us to examine effort levels directly.

Overall, the lab results are in line with our theory. In the two-person tournaments, effort is higher when contestants expect to be close in terms of luck (that is, if the density function is either increasing or decreasing – compared to effort in the uniform density case). Second, with larger groups, effort is substantially lower than in smaller groups if subjects expect not to receive a good draw (if the density function is decreasing), but they are essentially equal if subjects do expect to receive a good draw (in case of an increasing density function). Consistent with our theory under risk aversion, effort is smaller in large groups if the density function is uniform. The field data complement these insights by providing evidence consonant with the theory within a special case of the theory—when the density function is negatively skewed. In this case, we find evidence that adding competitors decreases individual effort levels especially when we control for fatigue.

We view our results as having import in several circles. For instance, they provide a theoretical basis for the disparate views concerning the optimal number of players in a contest, and clarify when larger tournaments should induce greater levels of effort. Such insights might aid the contest designer interested in optimal wage schemes, government procurement contracts for R&D contests, company promotional policies, and optimal mechanism design more generally.

Methodologically, this study showcases that by controlling the type of uncertainties characterizing the contest process, a crisp view of the impact of the number of contestants on a contestant's effort can be achieved. Likewise, by controlling for the number of competitors, one can estimate the effects of changing the nature of the uncertainty component. Gathering these insights across environments permits one to make much stronger inference than one could with either in isolation. This is so because our field experiment can check the robustness of laboratory results in a natural setting, where the mathematical assumptions of the theory cannot necessarily be guaranteed to hold. This approach provides a useful middle ground between the controlled environment of the laboratory and the unruly nature of uncontrolled field data.

References

Becker, Brian E., and Mark A. Huselid. 1992. "The Incentive Effects of Tournament Compensation Systems." *Administrative Science Quarterly,* 37: 336-50.

Boudreau, K., Lacetera, N., and Lakhani, K. 2011. "Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis." *Management Science*, 57(5): 843-863

Boudreau, K., Helfat, C., Lakhani, K., and Menietti, M. 2011. "Performance Responses to Competition Across Skill-Levels in Rank Order Tournaments: Field Evidence and Implications for Tournament Design." *HBS working paper14-014*

Bull, Clive, Andrew Schotter, and Keith Weigelt. 1987. "Tournament and Piece Rates: An Experimental Study." *Journal of Political Economy,* 95: 1-31.

Carmichael, Lorne. 1983. "Firm-specific Human Capital and Promotion Ladders." *Bell Journal of Economics,* 14: 251-8.

Chan, William, Pascal Courty, and Hao Li. 2009. "Suspense: Dynamic Incentives in Sports Contests." *Economic Journal,* 119: 24-46.

Che, Yeon-Koo, and Ian Gale. 2003. "Optimal Design of Research Contests." *American Economic Review,* 93(3): 646-71.

Eckel, Catherine, and Rick Wilson, 2004, "Is Trust a Risky Decision?" *Journal of Economic Behavior and Organization* 55, 447–65.

Ehrenberg, Ronald, and Michael Bognanno. 1990. "Do Tournaments Have Incentive Effects?" *Journal of Political Economy,* 98: 1307-24.

Eriksson, Tor. 1999. "Executive Compensation and Tournament Theory: Empirical Tests on Danish Data." *Journal of Labor Economics,* 17: 262-80.

Fischbacher, Urs. 2007. “Z-Tree: Zurich Toolbox for Ready-made Economic Experiments.” *Experimental Economics,* 10(2): 171-8.

Green, Jerry, and Nancy Stokey. 1983. “A Comparison of Tournaments and Contracts.” *Journal of Political Economy,* 91: 349-64.

Harbring, Christine and Bernd Irlenbusch. 2003. “An Experimental Study on Tournament Design.” *Labour Economics,* 10: 443-64.

Holmstrom, Bengt. 1982. “Moral Hazard in Teams.” *Bell Journal of Economics,* 13: 324-40.

Holt, Charles, and Susan Laury. 2002. “Risk Aversion and Incentive Effects.” *American Economic Review,* 92(5): 1644-55.

Lange, Andreas, John A. List, and Michael K. Price, 2007. “Using Lotteries to Finance Public Goods: Theory and Experimental Evidence,” *International Economic Review*, 48(3), pp. 901-927.

Lazear, Edward, and Sherwin Rosen. 1981. “Rank-order Tournaments as Optimum Labor Contracts.” *Journal of Political Economy,* 89(5): 841-64.

Levitt, Steven D., and John A. List. 2007. “What do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?” *Journal of Economic Perspectives,* 21(2): 153-74.

Liao, James C., David N. Beal, George V. Lauder, and Michael S. Triantafyllou. 2003. “Fish Exploiting Vortices Use Less Muscle.” *Science,* 302: 1566-9.

Loury, Glenn. 1979. “Market Structure and Innovation.” *Quarterly Journal of Economics,* 93(3): 395-410.

Lynch, James G., and Jeffrey S. Zax. 1998. “Prizes, Selection and Performance in Arabian Horse Racing.” *Discussion Paper*, University of Colorado.

Lynch, James G., and Jeffrey S. Zax. 2000. “The Rewards to Running: Prize Structure and Performance in Professional Road Racing.” *Journal of Sports Economics,* 1: 323-40.

Main, Brian G.M., Charles A. O’Reilly, and James Wade. 1993. “Top Executive Pay Tournaments of Teamwork.” *Journal of Labor Economics,* 11: 606-28.

Malcomson, James. 1984. “Work Incentives, Hierarchy, and Internal Labor Market.” *Journal of Political Economy,* 92: 486-507.

Nalebuff, Barry and Joseph Stiglitz. 1983. "Prizes and Incentives: Toward a General Theory of Compensation and Competition." *Bell Journal of Economics,* 14(1): 21-43.

O'Keefe, Mary, W. Kip Viscusi, and Richard Zeckhauser. 1984. "Economic Contests: Comparative Revenue Schemes." *Journal of Labor Economics,* 2: 27-56.

O'Reilly III, Charles A., Brian G. Main, and Graef S. Crystal. 1988. "CEO Compensation as Tournament and Social Comparison: A Tale of Two Theories." *Administrative Science Quarterly,* 33(2): 257-74.

Orrison, Alannah, Andrew Schotter, and Keith Weigelt. 2004. "Multiperson Tournaments: An Experimental Examination." *Management Science,* 50(2): 268-79.

Orszag, Jonathan M. 1994. "A New Look at Incentive Effects and Golf Tournaments." *Economics Letters,* 46: 77– 88.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms", *Journal of Economic Literature,* 37: 7-63.

Schumpeter, Joseph. 1950. *Capitalism, Socialism, and Democracy*, Third edition, New York: Harper &Row.

Simpson, R. David, Roger Sedjo, and John Reid. 1996. "Valuing Biodiversity for Use in Pharmaceutical Research." *Journal of Political Economy,* 104(1): 163-85.

Zhou, Haiwen. 2002. *Three Essays on Industrial Organization*, Ph. D. Dissertation, University of Maryland, College Park.

Table 1 – Predicted equilibrium effort levels in each of the 6 treatments assuming that either all agents are risk neutral, or risk averse.

	Density		
	Decreasing f'<0	Uniform f'=0	Increasing f'>0
2-player contest	D2	U2	I2
Prediction Risk Neutral	33.33	25.0	33.33
Prediction Risk Averse	31.28	23.59	31.28
4-player contest	D4	U4	I4
Prediction Risk Neutral	22.86	25.0	42.86
Prediction Risk Averse	19.35	21.13	35.50

Notes: Treatment U4 denotes a uniform density 4-player contest. Risk neutral (risk averse) subjects are predicted to choose effort level 25.0 (21.13) in U4 when groups are homogenous in terms of risk preferences.

Table 2 – Lab results (average effort; standard deviations in parentheses).[a]

	Density		
	Decreasing	Uniform	Increasing
Treatment	$f' < 0$	$f' = 0$	$f' > 0$
2-player contest	D2	U2	I2
	44.202	31.920	35.468
	(14.989)	(15.628)	(16.022)
	# *subj.* = 48	# *subj.* = 40	# *subj = 48*
4-player contest	D4	U2	I4
	26.960	29.435	35.079
	(12.022)	(11.550)	(9.575)
	# *subj = 48*	# *subj.* = 40	# *subj = 48*
Comparison X2 vs X4, X = {D,U,I} (*p*-values, MW)	0.0013	0.462	0.763

[a] p-values obtained from Mann-Whitney U tests using effort, averaged over all periods and over all subjects in a group, as unit of observation ($N = 24$ in D2 and I2, $N = 20$ in U2; $N = 12$ in D4 and I4, and $N = 10$ in U4).

Table 3 - Results of the OLS regression explaining effort using all observations (column i) and using observations from period 5-10 (column ii).

	(i)	(ii)
Intercept	32.54***	36.24***
	(2.309)	(3.947)
D2	12.20***	14.27***
	(1.737)	(2.260)
I2	3.569**	5.185**
	(1.713)	(2.224)
D4	-5.039***	-5.852**
	(1.763)	(2.276)
U4	-2.485	-2.529
	(1.901)	(2.490)
I4	3.179*	4.814**
	(1.889)	(2.453)
Risk Attitude	-0.531**	-0.493
	(0.256)	(0.332)
t	-0.123	-0.666*
	(0.178)	(0.380)
Dummy "treatment played first	6.521***	5.610***
in a session"	(1.016)	(1.306)
N	2720	1632
F value	22.00	16.78

Robust standard errors are reported in parenthesis under the coefficient estimates. ***, **, * Significant at the 1%, 5%, and 10% respectively.

Table 4 – Field results (average of an entire tournament)

	Period				Change in group size (GS)		
	1	2	3	4	Wilcoxon Signed-Rank Test		
A. Sessions 1+3	2-player	8-player	2-player	8-player	GS increases (first time)	GS decreases (first time)	GS increases (second time)
Effort	0.723	0.646	0.641	0.672	p=0.020	p=0.963	p=0.507
St.dev.	(0.239)	(0.248)	(0.210)	(0.277)			
	Period						
	1	2	3	4	Wilcoxon Signed-Rank Test		
B. Sessions 2+4	8-player	2-player	8-player	2-player	GS decreases (first time)	GS increases (first time)	GS decreases (second time)
Effort	0.597	0.566	0.518	0.524	p=0.294	p=0.088	p=0.754
St.dev.	(0.229)	(0.189)	(0.211)	(0.235)			

Notes: Effort intensity is the average number of casts per minute, corrected for the time elapsed between fish caught and the moment at which a fisherman restarts fishing. Wilcoxon Signed-Rank Test is for within person change of effort, where n = 32 for sessions 1 + 3 and n = 32 for sessions 2 + 4.

Table 5 – OLS estimation results of percentage changes in effort levels in the field

Variable	(i) One hour average	(ii) Average over the first 15 minutes
Decreasegroupsize	0.0963*	0.149*
	(0.058)	(0.077)
Session 1	-0.0902**	-0.204**
	(0.043)	(0.099)
Session 2	-0.158***	-0.221***
	(0.058)	(0.075)
Session 3	-0.127**	-0.153**
	(0.058)	(0.069)
Session 4	-0.248***	-0.193**
	(0.089)	(0.077)
Number of observations	192	192
F-value	3.52	3.68

Notes: Dependent variable is a subject's percentage change in effort between periods. Decreasegroupsize is a dummy variable which has a value of 1 if group size decreases from $n = 8$ in period t-1 to $n = 2$ in period t. The session variables capture session specific fixed effects, and their coefficients measure the average rate of decline in effort over all four rounds. Robust standard errors are reported in parenthesis under the coefficient estimates. ***, **, * Significant at the 1%, 5%, and 10% respectively.

Appendix I: A Model to Explore the Effect of Effort on Catch

In this appendix, we present the catch and effort data of a pilot session. Sixteen fishermen were placed in four treatments. In the first two treatments, the fishermen could catch as much fish as they could for a period of 30 minutes, each fish yielding €1 and €5 respectively. Treatment three was a tournament of thirty minutes where the winner earns €2 for each fish caught. The final treatment was a tournament of thirty minutes where the winner's prize was €5. The table below shows the OLS results of a regression on effort and catch. The results show a high correlation between rod casts and catch, which legitimizes the use of this variable as a measure of effort.

Table A1 – OLS estimation results for catch of fish in the pilot study

Dependent Variable: catch of fish	
Effort intensity	2.430***
	(0.564)
N	64
R^2	0.218

Notes: Dependent variable is subject's catch of fish in a period. Effort intensity is the number of casts per minute, corrected for the time elapsed between fish caught and the moment at which a fisherman restarts fishing. *** Significant at the 1% level.

Similar results are obtained when using a negative binomial model rather than OLS. The data of the pilot study used to estimate the production function of catch can be used to study the shape of the density function. The figure below plots the frequency distribution of the error terms, indicating that "bad luck" (at the level of the individual) occurs much more frequently than "good luck".

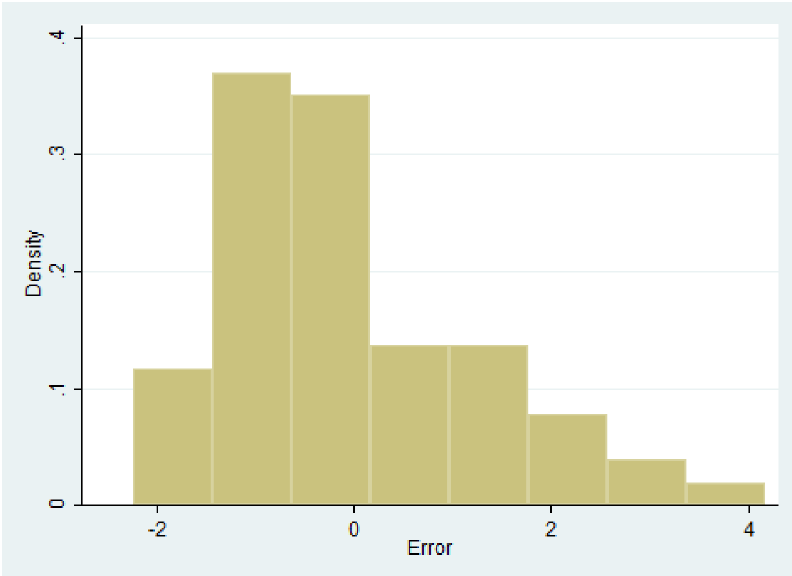


Figure A.1: The density of the random shock.

Appendix II. Experimental Instructions for D2D4 Treatment (to be made available online)

Introduction

This is an experiment in decision–making. You receive €4 for participating in this session. In addition, you can earn money because of the decisions you make in the experiment. All your earnings will be paid to you via bank transfer within the next 48 hours.

Before we start, we would like to ask that you do not communicate with other people during this session. Please also turn off your mobile phone.

The experiment consists of three parts: Part I, II and III. In all three parts you can earn points. You will be paid 1 euro for every point you have earned in the experiment. That is,

1 point is worth 1 euro, 5 points are worth 5 Euro's.

All your earnings in this experiment will be paid to you via bank transfer within 48 hours.

The instructions for the first part, Part I, will be read out aloud now, and you are invited to read along. After completion of Part I of the experiment, you will receive the instructions for Part II, and then for part III.

Part I of the Experiment

The first part of the experiment consists of 10 decision rounds. In each of the ten rounds you are requested to make one decision that we will explain in a moment. With this decision, you can earn points. However, in Part I only one of those rounds will be paid out. At the end of Part I, the computer will randomly select one of the ten decision rounds, and you will receive the points that you earned in that decision round.

Before the first round of Part I, you will be randomly grouped with **one** other participant in this room. That participant will be called your "group member". Your group member will remain the same individual throughout the entire first part of the experiment. The identity of your group member will not be revealed to you and your identity will not be revealed to him or her.

Experimental Procedure in each of the ten rounds of Part I

In every round of Part I, your Revenues are either equal to 5 points, or to 2 points. Whether you receive Revenues of 5 points or 2 points in a round, depends on whether or not you have the highest Total Number of your group in that round. A participant's Total Number in a round is the sum of the Decision Number that that participant chose in that round, and a Random Number that was assigned to that participant in that round.

Total Number = Decision Number + Random Number.

In every round of Part I, all participants first choose a Decision Number. After all participants have chosen their Decision Number, the computer randomly selects a Random Number for each participant. As stated before, a participant's Total Number is the sum of the Decision Number he/she chose, and the Random Number that was assigned to him/her. In every group, the participant with the highest Total Number receives Revenues of 5 points; the other participant receives Revenues of 2 points.

Your Total Earnings in a round are equal to your Revenues minus the Decision Costs you incur because of the Decision Number you chose. You can choose a Decision Number between 0 and 100. As shown in Table 1 (separate from this document, on your desk), the higher the number you choose, the higher the costs you incur. Costs are measured in points too. In the first column of the table you see the numbers 0 to 100. Associated with each Decision Number are Decision Costs, shown in the second column. For example, if you choose Decision Number = 20, your costs are 0.12 points (= 12 euro cents). All participants have the same "Decision Costs Table".

Your Total Earnings in a round are

5 – Decision Costs	if you have the highest Total Number of your group
2 – Decision Costs	if you do not have the highest Total Number of your group

If you and the other group member have the same Total Number, the computer will randomly decide which of the two receives 5 points.

The Random Number

Both you and the other member in your group choose a Decision Number. You and the other group member make their decisions independently and at the same time. After each group member has chosen his/her Decision Number, a Random Number is generated for each participant by the computer. Random Numbers are drawn from the range between – 100 and + 100.

When you submit your Decision Number, you do not know what Random Number you will receive, and also not what Random Number the other member of your group will receive. But before we proceed, consider the following hypothetical example.

Hypothetical example for Part I

Suppose that you WERE informed about the difference in Random Numbers received by you and by the other group member. The participant drawing the lowest Random Number (you, or your other group member) can then calculate how much larger his/her Decision Number needs to be (as compared to the Decision Number chosen by the participant receiving the highest Random Number) to receive Revenues of 5 points rather than Revenues of 2 points.

Suppose that the difference in Random Numbers is very large. If you received the highest Random Number, you know that you are likely to receive 5 points in this round even if you choose a low Decision Number. If you received the lowest Random Number, you know that you have to incur substantial Decision Costs to receive Revenues of 5 points in this round rather than Revenues of 2 points.

Suppose that the difference in Random Numbers is very small. Independent of whether you received the highest or the lowest Random Number, you know that it crucially depends on the Decision Numbers chosen by you and by the other group member whether you receive 5 points in this round, or 2 points. A small increase in the Decision Number chosen may result in you receiving Revenues of 5 points instead of 2 points.

This example is hypothetical, because NO participant is informed of the Random Number he/she receives until after he/she has chosen his/her Decision Number. So at the time you choose your Decision Number, you do NOT know with certainty whether the difference in Random Numbers will be large or small. But you can form expectations about whether the difference in Random Numbers is *likely* to be large, or small.

Let us now explain how Random Numbers are generated.

Figure 1 shows how likely it is that you receive a specific Random Number between – 100 and +100. On the horizontal axis the range of Random Numbers is presented, and on the vertical axis we show the probability that a participant receives that particular Random Number. For example, the probability that a participant receives a Random Number of, say, –50 is three times higher than the probability that he/she receives a Random Number of +50. The probability of receiving – 75 is six times higher than the probability of receiving +75.

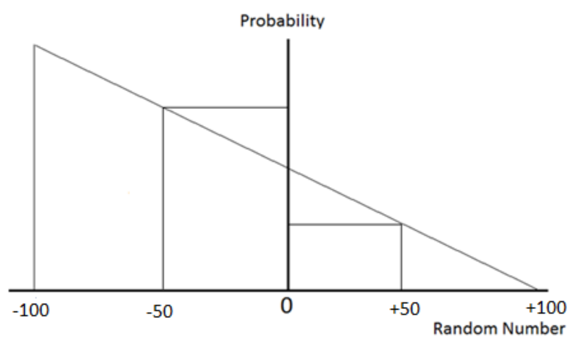


Figure 1

For each participant it is much more likely that he/she will receive a negative Random Number than a positive one.

This is also shown in Figure 2. In fact, the probability that an individual participant receives a Random Number between 0 and +100 is 25%, and hence the probability that he/she receives a Random Number between –100 and 0 is equal to 75%. There is one other participant in your group. That means that the probability that both you and the other participant receive a negative Random Number is slightly higher than 56% (0.75×0.75 = 0.5625). The probability that both you and the other participant receive a positive number is slightly higher than 6% (0.25×0.25 = 0.0625). Summing up, the probability that you and the other group member receive a similar Random Number, is higher than 50%. In fact, it is equal to 62.5% (that is, 0.5625 + 0.0625 = 0.625).

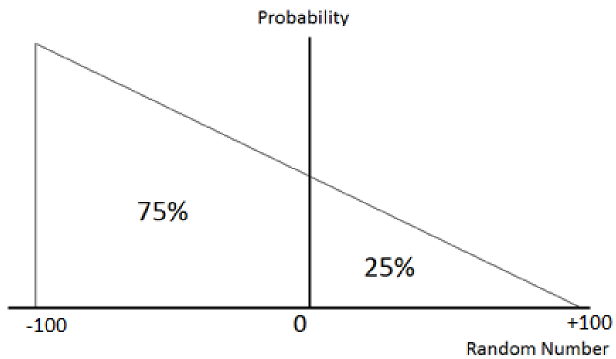


Figure 2

Having determined whether your Revenues are 5 points or 2 points, the computer next calculates your Total Earnings in the round by subtracting the Decision Costs you incur because of the Decision Number you chose from the Revenues you receive (5 points, or 2 points). After having been informed of your net earnings, the next round begins.

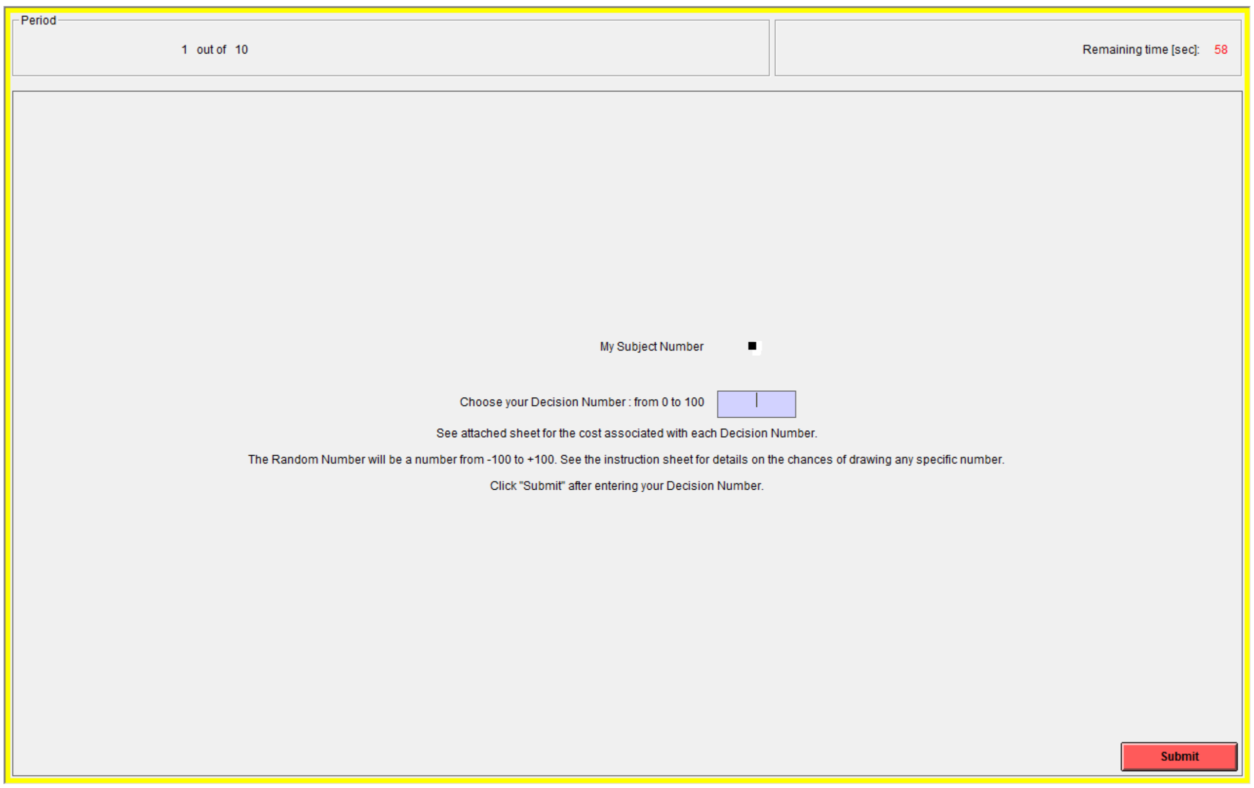
Rounds 2-10 of Part I

After Round 1 is completed, you will perform the same procedures for Round 2, Round 3, and so on for 10 rounds. In each round you will choose a Decision Number (of course, you may choose the same one in all rounds, or different ones in every round), the computer will again generate a Random Number when you submit your Decision Number, your Total Number will be compared to the Total Number of the other member of your group, and the computer will calculate your Total Earnings for the round.

When round 10 is completed, the computer randomly selects one of the ten rounds. This is the round that counts for your earnings in Part I.

Screens of Part I

Having completed the description of the task in Part I, let us now have a look at the screens. In the first screen you are requested to enter your Decision Number.

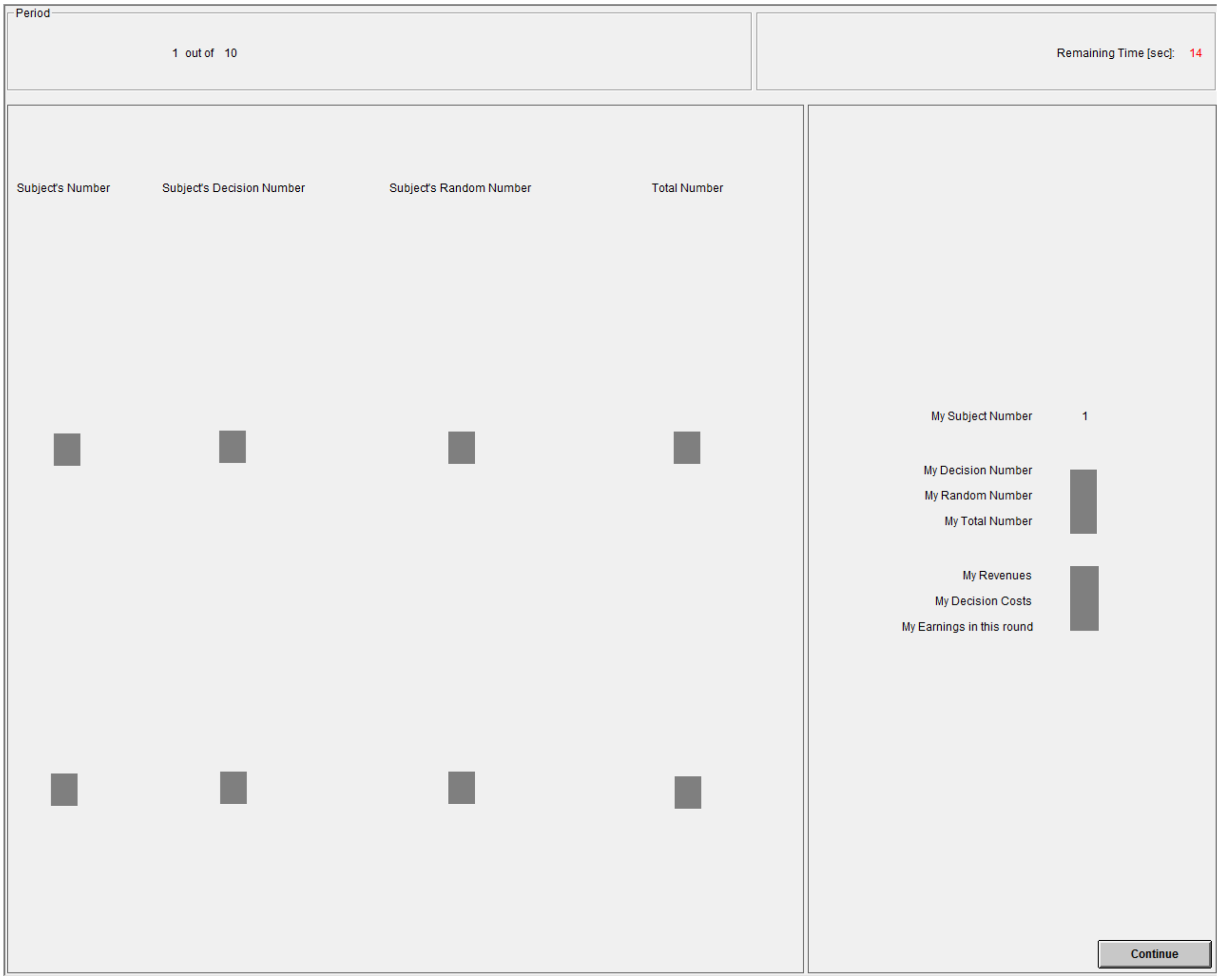


Screen 1

Type in your Decision Number and click "Submit".

After each participant has made his/her decision, the computer randomly generates a Random Number between –100 and +100, for each group member separately. The probabilities of receiving a specific Random Number are presented in Figure 1.

Then, Screen 2 appears. This screen shows your Decision Number as well as your Random Number. You will also receive feedback on the Decision Number and Random Number of your other group member. If your Total Number is higher than the Total Number of the other participant in your group, your Total Earnings in this round are 5 points minus the Decision Costs you incur because of the Decision Number you chose. If your Total Number is lower than the Total Number of the other participant in your group, your Total Earnings in this round are 2 points minus the Decision Costs you incur because of the Decision Number you chose. In case of ties in the Total Numbers, the computer will randomly decide who receives the 5 points Revenues, and who receives the 2 points Revenues.



Screen 2

At the end of each round, a third screen will appear that summarizes the history of your choices and the outcomes in all previous rounds.

This completes the description of the screens.

If there are no further questions, we now start with a short questionnaire. Please answer the following questions. When you have finished answering them, please raise your hand and we will come by to check your answers.

Test Questions Part I

1) With how many <u>other</u> participants do you form a group?___________________

2) Please circle the correct option. If my Revenues in a round are equal to 5 points, this means that in that round

   a. my Decision Number must have been higher than that of the other member of my group.

   b. my Random Number must have been higher than that of the other member of my group.

   c. the sum of the Decision Number and the Random Number of the other member of my group was not higher than the sum of my Decision Number and my Random Number.

3) If the Decision Number I choose is 100, then my Decision Costs are__________________ points.

   If the Decision Number I choose is 50, then my Decision Costs are__________________ points.

   If the Decision Number I choose is 50 and the Random Number I receive is 25, then my Decision Costs are___________________ points.

4) Please circle the correct option. The Random Number that is generated:

   a. is more likely to be negative (that is, the probability is larger than 50%),

b. is equally likely to be negative or positive,

c. is more likely to be positive (that is, the probability is larger than 50%).

5) The probability that both I and the other participant receive a similar Random Number (the probability that both a negative Random Number, PLUS the probability that both a positive Random Number), is _____%.

Part II of the experiment

The task in this second part of the experiment is the same as the previous part. As was the case in Part I, Part II will also consist of 10 rounds. Again, only one of these rounds will be paid, and at the end of Part II the computer randomly selects the round for which you will be paid.

As was in the case in Part I, you choose a Decision Number, after which a Random Number is assigned to you. The Decision Costs are the same as in Part I (see Table 1 on your desk), and Figure 1 of Part I also represents how likely it is that you receive a specific positive Random Number in Part II as well.

As before, your Total Earnings in a round in Part II are equal to

5 – Decision Costs	if you have the highest Total Number of your group, or
2 – Decision Costs	if you do not have the highest Total Number of your group

In case of ties in Total Numbers, the computer will randomly decide which group member receives Revenues of 5 points, and who receives Revenues of 2 points.

The ONLY difference between Part I and Part II is that before the first round of Part II, you will be randomly grouped with **three** other participants, whereas you were grouped with one other participant in Part I. At the beginning of Part II all participants are randomly regrouped in groups of four (you, and three other participants). That means that it is not very likely that the same individual you were matched with in Part I, is among the three group members you are matched with in Part II. The other three members of your group are the same individuals in each of the ten rounds in Part II.

The identity of the other members of your group will not be revealed to you and your identity will not be revealed to any of them.

As was the case in Part I**,** a Random Number is generated for you by the computer in every round of Part II after each group member has chosen his/her Decision Number. A Random Number is also generated separately for each of the other three members of your group.

When you submit your Decision Number, you do not know what Random Number you will receive, and what Random Number each of the other three group members will receive. But before we proceed, consider the following hypothetical example.

A hypothetical example for Part II

Suppose that you WERE informed about the difference in Random Numbers received by you and by each of the other three group members.

If you received a very high Random Number, choosing a slightly higher Decision Number may result in you receiving Revenues of 5 points instead of 2 points, but only if AT LEAST one of the three other group members also received a very positive number. (If no other participant would have received a very high Random Number, you would receive Revenues of 5 points anyway.)

If you received a very small Random Number, choosing a slightly higher Decision Number may result in you receiving Revenues of 5 points instead of 2 points, but only if NONE of the three other group members received a very positive Random Number either. (If one or more other participants would have received a very high Random Number, you would have to incur very high Decision Costs to receive Revenues of 5 points rather than of 2 points).

This example is hypothetical, because NO participant is informed of the Random Number he/she receives until after he/she has chosen his/her Decision Number. So at the time you choose your Decision Number, you do NOT know with certainty whether you will receive a high Random Numbers or a very low Random Number, or what Random Numbers are received by the other three group members. But you can form expectations about how likely it is that ALL four group members receive a very low Random Number, and how likely it is that you receive a very high Random Number and at least one other group member does so too.

As was the case in Part I, for each participant it is much more likely that he/she will receive a negative Random Number than a positive one, see Figure 2 of Part I. The probability that an individual participant receives a Random Number between 0 and +100 is 25%, and hence the probability that he/she receives a Random Number between –100 and 0 is equal to 75%. That means that the probability that all four group members receive a negative Random Number, is equal to 32% (that is, $(0.75)^4 = 0.32$). The probability that you receive a positive Random Number and at least one other group member as well, is equal to 14% (that is, $0.25 \times (1 - (0.75)^3) = 0.14$). The probability that either all four group members receive a negative Random Number or that you and at least one other participant receive a positive Random Number, is thus less than 50% (it is 0.32 + 0.14 = 0.46).

Rounds 2-10 in Part II

After Round 1 is completed, you will perform the same procedures for Round 2, and so on for 10 rounds. In each round you will choose a Decision Number (of course, you may choose the same one in all rounds, or different ones in every round), the computer will generate a new Random Number for you after you submitted your Decision Number, your Total Number will be compared to the Total Number of each of the three other members of your group, and the computer will calculate your earnings for the round.

When Round 10 is completed, the computer randomly selects one of the ten rounds. This is the round that counts for your earnings in Part II.

The screens of Part II

The screens of this second part of the experiment look very similar to the screens of Part I, except that you now receive information of the Decision Numbers and Random Numbers of all **three** other participants in your group.

Test questions for Part II

1) With how many <u>other</u> participants do you form a group in Part II?_____________

2) Please circle the correct option. The Random Number that is generated:

    a. is more likely to be negative (that is, the probability is larger than 50%),

    b. is equally likely to be negative or positive,

    c. is more likely to be positive (that is, the probability is larger than 50%).

3) The probability that (i) all four members of my group receive a negative Random Number, PLUS the probability that (ii) I receive a positive Random Number and at least one other participant in my group also receives a positive probability, is _____%.

Appendix III: Description of the Risk Preference Elicitation Task (to be made available online)

Part III of the Experiment

In this third part of the experiment, you will be making ten choices between two options, OPTION A and OPTION B. Each option is a lottery, and every combination of the lotteries in Option A and B, are called Choice Pairs. On the following Screen, the first column denotes the Choice Pairs, numbered from 1 to 10. The second column presents the details of the lottery of Option A, and the fourth presents the details of the lottery of Option B.

Choice Pair	Option A	Your Choice	Option B
1	€2 if 1 €1,60 if 2-10	A ○ ○ B	€3,85 if 1 €0,10 if 2-10
2	€2 if 1-2 €1,60 if 3-10	A ○ ○ B	€3,85 if 1-2 €0,10 if 3-10
3	€2 if 1-3 €1,60 if 4-10	A ○ ○ B	€3,85 if 1-3 €0,10 if 4-10
4	€2 if 1-4 €1,60 if 5-10	A ○ ○ B	€3,85 if 1-4 €0,10 if 5-10
5	€2 if 1-5 €1,60 if 6-10	A ○ ○ B	€3,85 if 1-5 €0,10 if 6-10
6	€2 if 1-6 €1,60 if 7-10	A ○ ○ B	€3,85 if 1-6 €0,10 if 7-10
7	€2 if 1-7 €1,60 if 8-10	A ○ ○ B	€3,85 if 1-7 €0,10 if 8-10
8	€2 if 1-8 €1,60 if 9-10	A ○ ○ B	€3,85 if 1-8 €0,10 if 9-10
9	€2 if 1-9 €1,60 if 10	A ○ ○ B	€3,85 if 1-9 €0,10 if 10
10	€2 if 1-10	A ○ ○ B	€3,85 if 1-10

FINISH

Please have a look at the lottery of Option A in Choice Pair 1. The computer randomly selects a number from the range 1, 2, 3, …10.. If the random number drawn is equal to “1”, this lottery pays €2.00; if the random number drawn is “2”, “3”, “4”, “5”, “6”, “7”, “8”, “9” or “10”, the lottery pays €1.60. Similarly, the lottery in Option B of Choice Pair 1 pays €3.85 if the randomly drawn number is equal to "1", and it pays €0.10 if the number drawn is "2", "3", "4", "5", "6", "7", "8", "9" or "10". In the third column you can indicate which of the two lotteries in Choice Pair 1 you prefer to participate in; the lottery as specified in Option A, or the lottery in Option B.

After you have indicated whether you prefer to participate in the lottery of Option A or in that of Option B in Choice Pair 1, move to the second Choice Pair, and indicate whether you prefer Option A or B in that second Choice Pair. In Choice Pair 2, the lottery in Option A pays €2.00 if the random number drawn is either "1" or "2", and it pays €1.60 in case the random number drawn is equal to 3, 4, 5, …, or 10. Similarly, the lottery in Option B of Choice Pair 2 pays €3.85 if the random number drawn is either "1" or "2", and it pays €0.10 in case the random number drawn is equal to 3, 4, 5, …, or 10. Again, you can indicate in the third column which of the two lotteries in Choice Pair 2 you prefer to participate in.

Note that the further down the screen you go, the larger the chances are of receiving the higher payoff in each of the two Options (€2.00 in Option A, and €3.85 in Option B), increases. In fact, in Choice Pair 10 you can receive €2.00 for certain if you choose Option A in that Choice Pair, or receive €3.85 with certainty if you choose Option B.

Please indicate for all ten Choice Pairs whether you prefer to participate in the lottery of Option A, or of B. Note that you can switch only once from Option A to Option B or vice versa, when you go down the list of Choice Pairs. Switching from Option A to Option B and then back to Option A, or vice versa, implies that at least one of your choices must be inconsistent.

While you make ten choices, only one of ten Choice Pairs will be used to determine your earnings. The computer randomly selects one of the Choice Pairs, looks up whether you chose Option A or Option B for that Choice Pair, randomly draws a number between 1 and 10, and determines how much money you receive as a result.

Your earnings will be shown on your computer screen after you have pressed the ‘Finish’ button. Earnings for this part of the experiment will be added to your previous earnings, and you will be paid all earnings by bank transfer within 48 hours.

Are there any questions? Please raise your hand, and we will come by to answer your question.

Appendix IV: The Experimental Instructions for the Field Experiment (Session 1 + 3)

IVA: Summary of rules handed out to the participants

Tournament

- You will participate in four tournaments. You will be assigned into groups which change in composition over the day. Therefore, it is likely that you participate in tournaments with changing participants.
- The duration of each tournament is 1 hour.
- The winner of a tournament is the one who catches most fish of his/her group.
- Rainbow trout and salmon trout both count as 1 fish.
- In case of a draw, the winner is the one who caught his/her first fish first.
- Whenever you catch a fish, make sure to communicate this to the organizers behind the desk. In that way, we can make sure that we do not make mistakes in counting the number of fish caught.
- For each tournament, only the winner receives a price. He/she receives €10.
- The beginning and end of a tournament is indicated by a blow on a whistle.
- Each tournament is a separate tournament with each its own winner. It does not matter who has the most fish at the end of the day.

Sequence of events

- Each participant plays 4 tournaments, tournament A through D.
- At the beginning of a new tournament you will change your fishing spot. Between the tournaments will be a break of 5 minutes.
- In tournament A, you will play in groups of 2. You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participants at spot

20 and spot 1 play a tournament, the participants at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.

- Tournament B is played in groups of 8. You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.
- Between tournament B and C is a break of 15 minutes.
- Tournament C is played in groups of 2 (just like tournament A). You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participants at spot 20 and spot 1 play a tournament, the participants at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.
- Tournament D is played in groups of 8 (just like tournament B). You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.

Putting fish into the pond

- For tournament A (which starts at 9.30 a.m.) we put 3 rainbow trout into the pond for each participant; 3 x 16 = 48 rainbow trout in total. In addition, we put in 10 extra rainbow trout. In total, 58 rainbow trout are put into the pond.
- For tournaments B, C, and D (which start at approximately 10.35 a.m., 11.50 a.m. and 12.55 p.m.) we put a number of rainbow trout into the pond equal to the total catch (both rainbow trout and salmon trout) of the previous tournament. This means that at the start of each tournament there is an equal number of fish in the pond.

Payment

- Your total earnings consist of your earnings in tournament A, B, C, and D.
- You are not allowed to keep each fish you catch! The total amount of caught fish is divided equally at the end of the day.
- For your participation you will receive €5.

IVB: Rules read out loud by the researcher

Welcome to this study by Tilburg University. Before we start, we want to point out two things. Firstly, this study is independent of the organization 'de Biestse Oevers'. We are grateful that we are allowed to conduct this study here, but this organization has nothing to do with what we are doing here. All responsibility lies with Tilburg University. Secondly, we want to make clear that this study has nothing to do with the well-being of animals, environmental causes or the like. As researchers, we accept the rules and habits of the sports fishing as it is practiced at 'de Biestse Oevers'. We cannot tell you the exact aim of this study. We do want to stress that your privacy is protected; none of the results we report can be traced on an individual level.

As you know, you don't have to pay to take part in this study. The fishing fee is paid by Tilburg University. Each fish you catch, you are allowed to take home. In addition, you can earn money.

We ask you to abide strictly by the rules which we impose.

The study

In the next four hours, we ask you to fish according to the rules as we will explain them now. All rules that normally hold at 'de Biestse Oevers' remain in place. This means that it is not permitted to throw fish you have caught back into the pond, you are only allowed to fish with one rod, you are only allowed to use a scoop net to set fish ashore, you are only allowed to use the usual types of bait, etc.

Today you will participate in four tournaments. Each tournament takes 1 hour. The winner of a tournament is the one who catches most fish of his/her group. You are allowed to catch as much fish as possible. The pond is mainly stocked with rainbow trout, but there may also be salmon trout in the pond. Each fish you catch carries equal weight in determining who wins a tournament.

In case of a draw between two or more participants, the winner is the one who caught his/her fish in the least amount of time. In case this also results in a draw, we will toss a coin to determine the winner.

Whenever you catch a fish, please communicate this to the organizers behind the table. Wait for them to answer your call (by means of a thumb raised in the air). In this way, we make sure that we do not make mistakes in counting the number of fish caught. For each tournament there is only a prize for the winner. He/she receives €10.

The beginning and end of each tournament is marked by a whistle; each tournament lasts exactly 1 hour. At the moment the second whistle sounds, you have to you’re your line and hook out of the water. If at that moment a fish is attached to your hook, you can land this fish and count it to your score.

The total duration of the study is about 4.5 hours, from 9.30 a.m. until 2.00 p.m. Each tournament is separate from the other tournaments. There is no prize for having caught the most fish at the end of the day.

You will play four tournaments. Two times you will participate in a tournament with 7 other participants (in a group of 8), and two times you will participate in a tournament with 1 other participant (in a group of 2). During the study, the composition of a group changes. The spot at which you fish is determined by means of a lottery. The first tournament, tournament A (starting at 9.30 a.m.) is played in groups of 2. You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participant at spot 20 and spot 1 play a tournament, the participant at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.

Tournament B (starting at 10.35 a.m.) is played in groups of 8 participants. You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.

Between tournament B and C there is a break of 15 minutes.

Tournament C (starting at 11.50 a.m.) is again played in groups of 2 (just like tournament A). You draw a fishing spot number out of a bag. The one who fishes opposite of you is the other participant of the tournament. The participant at spot 20 and spot 1 play a tournament, the participant at spot 19 and spot 2 play a tournament, and so on. The winner of each pair of participants is the one who catches most fish.

Tournament D is played in groups of 8 (just like tournament B). You draw a fishing spot number out of a bag. One group consists of the 8 spots on the canal side of the pond (these are numbers 1 through 4 and 17 through 20); the second group consists of the 8 spots on the meadow side of the pond (these are numbers 7 through 10 and 11 through 14). The winner of each of the two groups is the one of the 8 participants who catches most fish in this tournament.

Stocking fish

For the first tournament, tournament A (which starts at 9.30 a.m.) we put 3 rainbow trout into the pond for each participant; 3 x 16 = 48 rainbow trout in total. In addition, we put in 10 extra rainbow trout. In total, 58 rainbow trout are put into the pond.

For tournaments B, C, and D (which start at approximately 10.35 a.m., 11.50 a.m. and 12.55 p.m.) we put a number of rainbow trout into the pond equal to the total catch (both rainbow trout and salmon trout) of the previous tournament. This means that at the start of each tournament there is an equal number of fish in the pond.

Payment

You will receive 5 euro for your participation. In addition, you will receive 10 euro for each tournament which you have won.

You are not allowed to keep all fish that you have caught. All fish caught will be divided equally among all participants at the end of a tournament.

Questions

If you have any questions regarding the rules of this study, you can ask them now, but also during the study. We do not answer questions regarding how best to fish. We also do not answer questions regarding the nature of this study.